Journal of Differential Equations **178**, 541–573 (2002)
doi:10.1006/jdeq.2000.4008, available online at http://www.idealibrary.com on IDEAL®

Normal Forms for Nonautonomous Differential Equations

Stefan Siegmund[1]

Institut für Mathematik, Universität Augsburg, 86135 Augsburg, Germany
E-mail: stefan.siegmund@math.uni-augsburg.de

Received May 11, 2000; revised December 4, 2000

We extend Henry Poincarés normal form theory for autonomous differential equations $\dot{x} = f(x)$ to nonautonomous differential equations $\dot{x} = f(t, x)$. Poincarés nonresonance condition $\lambda_j - \sum_{i=1}^{n} \ell_i \lambda_i \neq 0$ for eigenvalues is generalized to the new nonresonance condition $\lambda_j \cap \sum_{i=1}^{n} \ell_i \lambda_i = \emptyset$ for spectral intervals.

1. INTRODUCTION

The famous French mathematician Henry Poincaré founded the normal form theory for autonomous differential equations $\dot{x} = f(x)$ near a rest point in his thesis in 1879. If the eigenvalues $\lambda_1, \ldots, \lambda_n$ of the linearization $\dot{x} = Df(x_0)\, x$ at the rest point x_0 satisfy the *nonresonance condition*

$$\lambda_j \neq \sum_{i=1}^{n} \ell_i \lambda_i\,, \tag{1}$$

$j \in \{1, \ldots, n\}$, $\ell_i \in \mathbb{N}_0 = \{0, 1, \ldots\}$, $\sum_{i=1}^{n} \ell_i \geqslant 2$, then the differential equation can be formally linearized. In this article Poincarés result is extended to nonautonomous differential equations $\dot{x} = f(t, x)$ near a reference solution. The normal form theory has applications in bifurcation theory (see e.g. Chow *et al.* [24], Crawford [29] and Kuznetsov [41]). Good presentations of the basics and introductions into normal form theory can be found, e.g., in V. I. Arnold [3, Chapter 5, pp. 170–190], Ashkenazi and Chow [5], Dora and Stolovitch [30] and Wiggins [65, pp. 211–252]. The

[1] This paper has been written while the author was a member of the "Graduiertenkolleg: Nichtlineare Probleme in Analysis, Geometrie und Physik" (GRK 283) financed by the Deutsche Forschungsgemeinschaft and the State of Bavaria.

0022-0396/02 $35.00

normal form theory is developed almost exclusively for autonomous differential equations and there exist only some results for special nonautonomous differential equations. In Goltser [33] the so-called Cesàro method of integration is used to calculate certain improper integrals by averaging. For that reason the results can be applied only to parameterdependent differential equations with an autonomous linear part and a nonautonomous nonlinearity of a special class (e.g. almost periodic). For hamiltonian systems with a small nonautonomous perturbation (especially for initial value problems of Schrödinger equations) Gompa [34] develops an "approximate normal form theory". Kostin [39] allows a nonautonomous, i.e., t-dependent, linear part assuming that its t-dependent eigenvalues satisfy certain asymptotic integrability conditions. Periodic differential equations are treated in Samovol [45] and Smith [55]. Well developed and elegantly notated is the normal form theory in L. Arnold [2] for random dynamical systems.

We extend some of the cited nonautonomous results and we generalize Poincarés crucial *nonresonance condition for the eigenvalues* of the linearized differential equation to a *new nonresonance condition for the spectral intervals* of the dichotomy spectrum (see Siegmund [52]).

The existing theory for autonomous differential equations or vector fields distinguishes between analytical, C^∞ and C^k normal forms. In a first step a formal transformation (i.e. a formal in general non-convergent series) is constructed with the Poincaré–Dulac scheme. Therefore one uses (in general infinitely many) polynomial transformations to eliminate all *nonresonant terms* of the vector field. To get an analytic normal form one has to answer the difficult question of convergence of the formal series. Extensive research of this convergence problem can be found, e.g., in Bruno[2] [16, 17] and [18, Part II, pp. 273–345]. For C^∞ normal forms see e.g. Belitskii [12]. It is often sufficient to consider normal forms up to a certain order $K \geqslant 1$. These so-called C^k normal forms, $k \leqslant K$ (for $k = K$ see also Belitskii [13]) are achieved by applying the Poincaré–Dulac scheme only finitely many times until terms of order higher than K do not change any more (i.e., one uses only a finite part of the formal transformation series). In a second step one constructs a C^k conjugacy which eliminates all Taylor coefficients of order K and higher order. Arguments to cut off such a Taylor rest were first used by Sternberg [56–59] and Chen [22]. By this procedure one gets a so-called resonant polynomial normal form (see, e.g., V. I. Arnold [4, Chapter 2.5, pp. 60–73] or Kopanskii [38]). The choice of the maximal k in dependence of K and the given differential equation is a difficult question and answers are known only for special cases (see, e.g., Belitskii [7], Kopanskii [37], Samovol [46, 47], Sell [49], and Stowe

[2] *Bruno* and *Brjuno* are different transcriptions of the same author.

[60]). In a final step one even tries to eliminate as many resonant terms as possible in the resonant polynomial normal form with a C^k conjugacy. Also this step is nontrivial and optimal results are known only for special situations, see, e.g., Bronstein and Kopanskii [19, Chapter II, pp. 29–192].

For further research and modern developments of normal form theory see also L. Arnold [2, Chapter 8, S. 405–463], Belitskii [8–11], Bonckaert and Dumortier [14], Bonckaert [15], Bronstein and Kopanskii [20], Chow *et al.* [25], ElBialy [31], Il'yashenko and Yakovenko [35], Katok and Hasselblatt [36, Chapter 2, pp. 57-104], Robinson [42], Sakamoto [44], Sell [48, 50], Takens [61, 62] and Warner and Sethna [64]. More bibliographical information can be found, e.g., in Chow and Hale [23, Chapter 12.8, pp. 449–450], Chow *et al.* [24, 2.12 Bibliographical Notes, pp. 188–190], and van der Meer [63, pp. 42–45].

Instead of giving an introduction into Poincarés normal form theory (which can be found in the cited references) we consider an example

$$\boxed{\begin{aligned} \dot{x} &= x \\ \dot{y} &= \lambda y + x^2 \end{aligned}}$$

with $\lambda \in \mathbb{R}$. We are looking for a near-identity transformation

$$H(x, y) = \begin{pmatrix} x \\ y \end{pmatrix} + h_2(x, y)$$

which eliminates the second-order nonlinearity $\binom{0}{x^2}$ and we choose $h_2 \in \operatorname{span}\left\{\binom{x^2}{0}, \binom{0}{x^2}, \binom{xy}{0}, \binom{0}{xy}, \binom{y^2}{0}, \binom{0}{y^2}\right\}$. It is not difficult to show that the transformed equation has no second-order nonlinearity if and only if the so-called *homological equation*

$$Ah_2(x, y) - Dh_2(x, y)\, A \begin{pmatrix} x \\ y \end{pmatrix} = f_2(x, y)$$

is satisfied with

$$A = \begin{pmatrix} 1 & 0 \\ 0 & \lambda \end{pmatrix} \quad \text{and} \quad f_2(x, y) = \begin{pmatrix} 0 \\ x^2 \end{pmatrix}.$$

It is solvable if and only if $\lambda \neq 2$ and with its unique solution we get

$$H(x, y) = \begin{pmatrix} x \\ y + \dfrac{1}{\lambda - 2} x^2 \end{pmatrix}.$$

For $\lambda \neq 2$ there exists no C^2 linearization of the differential equation (Dora and Stolovitch [30, Example 2.3, Proposition 2.1, pp. 149–151]). In this simple example the transformed equation $\dot{x} = x$, $\dot{y} = \lambda y$ is linear.

In general the elimination of second order nonlinearities produces higher order nonlinearities and the process has to be iterated. The resulting transformation is the composition of the transformations of each elimination step and it is *nonlinear* but is constructed by solving a *sequence of linear* equations.

In the following we consider a nonautonomous differential equation

$$\boxed{\dot{x} = f(t, x),} \tag{2}$$

$f: D \to \mathbb{R}^N$, $D \subset \mathbb{R} \times \mathbb{R}^N$ open, not in the vicinity of a rest point as Poincaré did, but in the vicinity of an arbitrary reference solution $\mu_0: \mathbb{R} \to \mathbb{R}^N$. Modern problems demand the use of nonautonomous differential equations with a right hand side which is only measurable in t. Constantin Carathéodory was the first who investigated this type of differential equations. In his Vorlesungen über relle Funktionen [21] he rigorously introduced the Lebesgue integral into the theory of ordinary differential equations. Today they bear his name. A nonautonomous differential equation (2) is of C^k Carathéodory type (see also Definition 2.1) if f is measurable in t (for fixed x) and C^k in x (for fixed t). For example, random differential equations $\dot{x} = f(\theta_t \omega, x)$ driven by a metric dynamical system θ are pathwise of Carathéodory type (L. Arnold [2]). Bilinear control systems $\dot{x} = [A_0 + \sum_{i=1}^{M} u_i(t)\, A_i]\, x$ for fixed controls $u: \mathbb{R} \to U$, $U \subset \mathbb{R}^M$ compact and convex, are of Carathéodory type (Colonius and Kliemann [26]). A function $\mu: I \to \mathbb{R}^N$ defined on some interval $I \subset \mathbb{R}$ is a (mild) solution of (2) if $\mu(t) - \mu(s) = \int_s^t f(u, \mu(u))\, du$ for $t, s \in I$. Such a solution μ is absolutely continuous, which is more than continuous and less then differentiable; indeed it is differentiable almost everywhere, see Craven [28, Prop. 5.2.10, 5.4.5, and 5.5.3]. If f is continuous then μ is a classical differentiable solution.

We will extend Poincarés normal form theory to differential equations of C^k Carathéodory type, $k \geqslant 2$, by showing that if the linearization $\dot{x} = D_x f(t, \mu_0(t))\, x$ of (2) along the reference solution μ_0 satisfies a nonresonance condition (see the Normal Form Theorem below), then the system (2) is locally C^k equivalent (see Definition 2.2) to a system $\dot{x} = g(t, x)$ in normal form, i.e., with zero reference solution, block diagonal linear part $\dot{x} = D_x g(t, 0)\, x$ and all nonresonant Taylor coefficients of g up to order k are zero.

We therefore have to use a proper replacement of the "linear algebra" for autonomous systems (i.e., eigenvalues and eigenspaces) in our nonautonomous situation. A spectral theory for nonautonomous differential equations is developed in Siegmund [52]. The dichotomy spectrum of the linearized differential equation $\dot{x} = D_x f(t, \mu_0(t))\, x$ consists of at most N closed disjoint intervals, in general the spectrum may be empty or unbounded. It consists of n, $1 \leqslant n \leqslant N$, compact disjoint intervals $\lambda_i = [a_i, b_i]$ if the system has bounded growth, i.e., if the evolution operator $\Phi(t, s)$ satisfies the estimate $\|\Phi(t, s)\| \leqslant Ke^{a|t-s|}$ for $t, s \in \mathbb{R}$ with constants $K \geqslant 1$, $a \geqslant 0$. If $\operatorname{ess\,sup}_{t \in \mathbb{R}} \|D_x f(t, \mu_0(t))\| < \infty$ then the linear system has bounded growth. More general sufficient conditions for bounded growth can be found in Coppel [27, p. 9]. For simplicity we assume that the linearized equation has bounded growth although the theory could also be developed in the general case.

2. PRELIMINARIES

DEFINITION 2.1. Let $N, M \in \mathbb{N} = \{1, 2, \ldots\}$, $D \subset \mathbb{R} \times \mathbb{R}^N$ open and $f\colon D \to \mathbb{R}^M$ a function.

(A) Then f is a *Carathéodory function* if for every interval $I \subset \mathbb{R}$ and every open set $U \subset \mathbb{R}^N$ with $I \times U \subset D$ the following holds:

(i) for a.a. $t \in I$ the mapping $f|_{I \times U}(t, \cdot) : U \to \mathbb{R}^M$ is continuous,

(ii) for all $x \in U$ the mapping $f|_{I \times U}(\cdot, x) : I \to \mathbb{R}^M$ is measurable (with respect to the Borel σ-algebras on I and $\mathbb{R}^M$).

(B) Then f is a C^k *Carathéodory function*, $k \geqslant 0$, if

(i) for a.a. $t \in \mathbb{R}$ and all $x \in \mathbb{R}^N$ with $(t, x) \in D$ the kth partial derivative $D_x^k f(t, x)$ exists,

(ii) for every $j \in \{0, \ldots, k\}$ the mapping $D_x^j f\colon D \to L^j(\mathbb{R}^N; \mathbb{R}^M)$ is a Carathéodory function.

THEOREM 2.1 (EXISTENCE OF SOLUTIONS). *Let $f\colon D \subset \mathbb{R} \times \mathbb{R}^N \to \mathbb{R}^N$ be a Carathéodory function with the property that for every $(\tau_0, \xi_0) \in D$ there exists an interval $I \subset \mathbb{R}$, a neighbourhood $U \subset \mathbb{R}^N$ and locally integrable functions $\ell_0, \ell_1\colon I \to \mathbb{R}_0^+$ with $(\tau_0, \xi_0) \in I \times U \subset D$,*

$$\|f(t, x) - f(t, \bar{x})\| \leqslant \ell_1(t)\, \|x - \bar{x}\| \qquad \textit{and} \qquad \|f(t, \xi_0)\| \leqslant \ell_0(t)$$

for all $x, \bar{x} \in U$ and a.a. $t \in I$. Then for every $(\tau_0, \xi_0) \in D$

$$\boxed{\dot{x} = f(t, x)}$$

has a unique (maximal) solution $\varphi(\cdot\,;\tau_0,\xi_0)\colon I_{\tau_0,\xi_0}\to\mathbb{R}^N$ *with* $x(\tau_0)=\xi_0$. *The set* $\Omega:=\{(t,\tau,\xi)\in\mathbb{R}^{1+1+N}:(\tau,\xi)\in D,\ t\in I_{\tau,\xi}\}$ *of definition is open and* $\varphi\colon\Omega\to\mathbb{R}^N$ *is continuous.*

Proof. We cite Kurzweil [40]. Using Definition 18.4.1, p. 331, and Theorem 18.4.13, p. 337, one gets uniqueness of solutions. Because of Remark 18.4.15, p. 338, it suffices to cite Theorem 12.1.1, p. 215, and Theorem 12.2.6, p. 223, for the remaining claims. ∎

There is no straightforward way to define a notion of conjugacy for nonautonomous differential equations. What do we mean by this? Two autonomous differential equations $\dot{x}=f_1(x)$ and $\dot{x}=f_2(x)$ in $\mathbb{R}^N$ are said to be conjugate if there exists a homeomorphism $H\colon\mathbb{R}^N\to\mathbb{R}^N$ such that the flows $\varphi_1(\cdot\,;\xi)$ resp. $\varphi_2(\cdot\,;\eta)$ satisfy the conjugacy relation $H(\varphi_1(t;\xi))=\varphi_2(t;H(\xi))$ for all $\xi\in\mathbb{R}^N$, $t\in I_\xi$, i.e. H maps solutions of the first equation onto solutions of the second equation and vice versa for H^{-1}. Now if we would define a conjugacy between two differential equations $\dot{x}=f_1(t,x)$ and $\dot{x}=f_2(t,x)$ of Carathéorody type by the same property, but now with a t-dependent H then for every $\tau\in I_{t,x}$

$$H(t,x):=\varphi_2(t;\tau,\varphi_1(\tau;t,x))$$

would establish a conjugacy, i.e., H maps solutions of the first equation onto solutions of the second equation and vice versa with

$$H^{-1}(t,x):=\varphi_1(t;\tau,\varphi_2(\tau;t,x)).$$

So in the nonautonomous situation we need some additional conditions which ensure that qualitative behaviour—at least for a single reference solution—is preserved under the transformation.

It is easy to see in the autonomous situation that for a conjugacy periodic solutions, limit sets and invariant sets of the first equation are bijectively mapped onto periodic solutions, limit sets and invariant sets, respectively, of the second equation and that (asymptotic) stability, attractivity and instability of bounded solutions is preserved under the conjugacy. In most cases this is enough, but note that the assumption of boundedness of solutions is essential for the preservation of stability. For example, the two linear systems $\dot{x}=1$, $\dot{y}=-y$ and $\dot{x}=1$, $\dot{y}=y$ are conjugate via $H(x,y)=(x,ye^{2x})$ but the first system is stable and the second is unstable. To preserve the stability of an unbounded solution μ it would be necessary to pose some uniformity condition on H, e.g., $\lim_{x\to0}H(\mu(t)+x)=H(\mu(t))$ uniformly in $t\in\mathbb{R}$. Such a uniformity condition is exactly what we need in the nonautonomous situation to define a meaningful notion of C^k equivalence (see Definition 2.2).

Define for $\varepsilon > 0$, $x_0 \in \mathbb{R}^N$, $\mu: \mathbb{R} \to \mathbb{R}^N$ the neighbourhoods

$$B_\varepsilon(x_0) = \{x \in \mathbb{R}^N : \|x - x_0\| < \varepsilon\}$$
$$U_\varepsilon(\mu) = \{(t, x) \in \mathbb{R} \times \mathbb{R}^N : x \in B_\varepsilon(\mu(t))\}.$$

Consider differential equations of C^k Carathéodory type, $k \geqslant 0$, with reference solutions

$$\dot{x} = f(t, x), \qquad \mu_0: \mathbb{R} \to \mathbb{R}^N \tag{3}$$
$$\dot{x} = g(t, x), \qquad \nu_0: \mathbb{R} \to \mathbb{R}^N \tag{4}$$

i.e., $f: D_f \subset \mathbb{R} \times \mathbb{R}^N \to \mathbb{R}^N$ and $g: D_g \subset \mathbb{R} \times \mathbb{R}^N \to \mathbb{R}^N$ are C^k Carathéodory functions. We assume that tubular neighbourhoods of the reference solutions are contained in the corresponding sets of definition; i.e., there exist $r > 0$ and $p > 0$ such that

$$U_r(\mu_0) \subset D_f \qquad \text{and} \qquad U_p(\nu_0) \subset D_g\,.$$

DEFINITION 2.2. Consider the two equations (3) and (4). If there exist r', p' with $0 < r' \leqslant r$ and $0 < p' \leqslant p$ together with continuous functions

$$H: U_{r'}(\mu_0) \to \mathbb{R}^N, \qquad H^{-1}: U_{p'}(\nu_0) \to \mathbb{R}^N,$$

then H is called a *local C^k equivalence between the system* (3) *with solution* μ_0 *and system* (4) *with solution* ν_0, if the following statements are valid:

(A) For each $t \in \mathbb{R}$ the mappings

$$H(t, \cdot) : B_{r'}(\mu_0(t)) \to H(t, B_{r'}(\mu_0(t))) \subset B_p(\nu_0(t))$$
$$H^{-1}(t, \cdot) : B_{p'}(\nu_0(t)) \to H^{-1}(t, B_{p'}(\nu_0(t))) \subset B_r(\mu_0(t))$$

are C^k diffeomorphisms (or homeomorphisms if $k = 0$) with

$$H(t, H^{-1}(t, x)) = x, \qquad \text{for} \quad x \in B_{p'}(\nu_0(t)),$$
$$H^{-1}(t, H(t, x)) = x, \qquad \text{for} \quad x \in B_{r'}(\mu_0(t)).$$

(B) If μ is a solution of (3) in $U_{r'}(\mu_0)$ then $H(\cdot, \mu(\cdot))$ is a solution of (4).
If ν is a solution of (4) in $U_{p'}(\nu_0)$ then $H^{-1}(\cdot, \nu(\cdot))$ is a solution of (3).

(C) The reference solutions are mapped uniformly onto each other:

$$\lim_{x \to 0} H(t, \mu_0(t)+x) = \nu_0(t) \quad \text{uniformly in} \quad t \in \mathbb{R},$$

$$\lim_{x \to 0} H^{-1}(t, \nu_0(t)+x) = \mu_0(t) \quad \text{uniformly in} \quad t \in \mathbb{R}.$$

LEMMA 2.1. *Consider the two equations* (3) *and* (4) *together with a solution* $\mu: I \to \mathbb{R}^N$ *of* (3) *which is defined on some interval* $I \subset \mathbb{R}$. *Then a mapping* $\nu: J \to \mathbb{R}^N$ *defined on some interval* $J \subset I$ *is a solution of*

$$\boxed{\dot{x} = g(t, x+\mu(t)) - f(t, \mu(t))} \tag{5}$$

if and only if $\nu + \mu: J \to \mathbb{R}^N$ *is a solution of the differential equation* (4).

Proof. Since μ is a solution of (3) one has for $t, s \in J$

$$\nu(t) - \nu(s) = \int_s^t g(\tau, \nu(\tau)+\mu(\tau)) - f(\tau, \mu(\tau))\, d\tau$$

$$\Leftrightarrow \nu(t) - \nu(s) + \int_s^t f(\tau, \mu(\tau))\, d\tau = \int_s^t g(\tau, \nu(\tau)+\mu(\tau))\, d\tau$$

$$\Leftrightarrow [\mu(t)+\nu(t)] - [\mu(s)+\nu(s)] = \int_s^t g(\tau, \nu(\tau)+\mu(\tau))\, d\tau$$

and the claim is proved. ▮

3. NORMAL FORMS

We consider a differential equation of C^k Carathéodory type, $k \geqslant 2$, satisfying the assumptions of Theorem 2.1 together with a reference solution

$$\dot{x} = f(t, x), \qquad \mu_0: \mathbb{R} \to \mathbb{R}^N, \tag{6}$$

and we additionally assume the following conditions:

- Set of definition: $r := \sup \{r' \geqslant 0 : U_{r'}(\mu_0) \subset D_f\} > 0$.
- Linearity: $\dot{x} = D_x f(t, \mu_0(t))\, x$ has bounded growth.
- Nonlinearity: $\|D_x^j f(t, \mu_0(t))\| \leqslant M$ for $2 \leqslant j \leqslant k$ and a.a. $t \in \mathbb{R}$.

We will simplify system (6) in three steps.

Step 1: *Trivialization of the Reference Solution*

Recall Lemma 2.1. If $f \equiv g$ then system (5) reduces to

$$\boxed{\dot{x} = f(t, x+\mu_0(t)) - f(t, \mu_0(t))} \tag{7}$$

which is usually called the differential equation of perturbed motion of (6) w.r.t. the solution μ_0. Obviously (7) has the zero solution and $v\colon J \subset I \to \mathbb{R}^N$ is a solution of (7) if and only if $v+\mu_0$ is a solution of (6). Because of Lemma 2.1 the mappings

$$\begin{aligned} R\colon U_r(\mu_0) &\to \mathbb{R}^N, \quad (t, x) \mapsto x - \mu_0(t) \\ R^{-1}\colon U_r(0) &\to \mathbb{R}^N, \quad (t, x) \mapsto x + \mu_0(t) \end{aligned}$$

define a C^∞ equivalence between (6) with reference solution μ_0 and the system (7) with zero reference solution. We rewrite (7) as

$$\boxed{\dot{x} = A_*(t)\, x + F_*(t, x),} \tag{8}$$

where $A_*(t) = D_x f(t, \mu_0(t))$ is the linear part and $F_*(t, x) = f(t, x+\mu_0(t)) - f(t, \mu_0(t)) - D_x f(t, \mu_0(t))\, x$ is the nonlinearity. Obviously $U_r(0) = \mathbb{R} \times B_r(0)$ is contained in the set of definition of the right-hand side of (8). Note that this simple transformation is a powerful nonautonomous tool. It is of no use in a purely autonomous framework.

Step 2: *Reduction of the Linear Part*

The Reduction Theorem in Siegmund [53] shows that there exists a kinematic similarity $S\colon \mathbb{R} \to \mathbb{R}^{N \times N}$ between the linearization $\dot{x} = A_*(t)\, x$ of (8) and a linear system

$$\dot{x} = A(t)\, x \tag{9}$$

such that $A\colon \mathbb{R} \to \mathbb{R}^{N \times N}$ is locally integrable and in block diagonal form

$$A(t) = \begin{pmatrix} A_1(t) & & \\ & \ddots & \\ & & A_n(t) \end{pmatrix}$$

and each block $A_i\colon \mathbb{R} \to \mathbb{R}^{N_i \times N_i}$, $i = 1, \ldots, n$, corresponds to a spectral interval λ_i. The dichotomy spectra $\Sigma(A_*)$ and $\Sigma(A)$ are the same and (9) has bounded growth.

LEMMA 3.1. *There exist p, p', r' with $0 < p' \leqslant p$ and $0 < r' \leqslant r$ such that the mappings*

$$U_{r'}(0) \to B_p(0), \qquad (t, x) \mapsto S^{-1}(t)\, x$$
$$U_{p'}(0) \to B_r(0), \qquad (t, x) \mapsto S(t)\, x$$

define a C^∞ equivalence between (8) *and the differential equation*

$$\boxed{\dot{x} = A(t)\, x + F(t, x)} \tag{10}$$

which is of C^k Carathéodory type with $F(t, x) = S(t)^{-1} F_*(t, S(t)\, x)$ and $U_p(0) \subset D_F$. Moreover $\|D_x^j F(t, 0)\| \leqslant M'$ for a.a. $t \in \mathbb{R}$ and all $j \in \{2, ..., k\}$ with some $M' > 0$.

Proof. Due to Coppel [27, p. 38] the kinematic similarity satisfies

$$\frac{d}{dt} S(t)^{-1} = A(t)\, S(t)^{-1} - S(t)^{-1} A_*(t).$$

Let μ be a solution of (8). Then $v(t) := S(t)^{-1} \mu(t)$ satisfies for a.a. $t \in \mathbb{R}$

$$\begin{aligned} \frac{d}{dt} v(t) &= \frac{d}{dt} S(t)^{-1}\, \mu(t) + S(t)^{-1} \frac{d}{dt} \mu(t) \\ &= [A(t)\, S(t)^{-1} - S(t)^{-1} A_*(t)]\, \mu(t) + S(t)^{-1} [A_*(t)\, \mu(t) + F_*(t, \mu(t))] \\ &= A(t)\, v(t) + S(t)^{-1} F_*(t, S(t)\, v(t)), \end{aligned}$$

i.e., v is a solution of (10) with $F(t, x) = S(t)^{-1} F_*(t, S(t)\, x)$. The remaining claims of the lemma can be easily shown. ∎

Step 3: *Elimination of Nonresonant Taylor Components*

This is the crucial step. We will eliminate Taylor components of the nonlinearity which correspond to the blocks $A_i(t) \in \mathbb{R}^{N_i \times N_i}$, $i = 1, ..., n$, of the linear part $A(t)$. We define $E_i := \mathbb{R}^{N_i}$, $i = 1, ..., n$, and write $F = (F_1, ..., F_n)$ with the component functions $F_i\colon D_F \to E_i$.

In order to present the ideas we first motivate the construction of the transformation and the nonresonance condition.

For simplicity assume therefore that the system (10) is globally defined, i.e., $D_F = \mathbb{R} \times \mathbb{R}^N$ and that each solution exists on $\mathbb{R}$, this can be achieved by cutting of F outside the neighbourhood $U_\varepsilon(0)$ of the zero solution. Now for a.a. $t \in \mathbb{R}$ we can expand $F(t, \cdot)$ into a Taylor series at $x = 0$

$$F(t, x) = \sum_{\ell \in \mathbb{N}_0^n : 2 \leqslant |\ell| \leqslant k} \frac{1}{\ell!} D_x^\ell F(t, 0) \cdot x^\ell + o(\|x\|^k).$$

Now we are looking for a condition under which a C^k transformation exists which eliminates the jth component $\frac{1}{\ell!} D_x^\ell F_j(t, 0) \cdot x^\ell$ of a summand in the Taylor expansion. Therefore choose and fix a $j \in \{1, ..., n\}$ and a multi index $\ell \in \mathbb{N}_0^n$ with $2 \leqslant |\ell| \leqslant k$. For simplicity we assume that the Taylor coefficients of F at $x = 0$ up to order $|\ell| - 1$ are already eliminated, i.e.,

$$D_x^\kappa F(t, 0) = 0 \quad \text{for a.a. } t \in \mathbb{R} \text{ and all } \kappa \in \mathbb{N}_0^n \text{ with } 2 \leqslant |\kappa| \leqslant |\ell| - 1 . \tag{11}$$

We define a new C^k Carathéodory function G: $\mathbb{R} \times \mathbb{R}^N \to \mathbb{R}^N$ by

$$G(t, x) := F(t, x) - \left(0, ..., 0, \frac{1}{\ell!} D_x^\ell F_j(t, 0) \cdot x^\ell, 0, ..., 0\right).$$

To derive some necessary conditions for the existence of an equivalence we assume now that a near identity C^k equivalence $H(t, x) = x + h(t, x)$ between (10) with zero reference solution and the differential equation

$$\boxed{\dot{x} = A(t)\, x + G(t, x)} \tag{12}$$

with zero reference solution exists, where h: $\mathbb{R} \times \mathbb{R}^N \to \mathbb{R}^N$ is a mapping with $h(t, 0) \equiv 0$ and $D_x h(t, 0) \equiv 0$ on $\mathbb{R}$. We will make some observations which will help us to construct an explicit candidate for a C^k equivalence.

First we will assign a differential equation to the values of the transformation H along a fixed solution.

Observation 1. For each initial condition $(\tau, \xi) \in \mathbb{R} \times \mathbb{R}^N$ the mapping $h(t, \varphi(t; \tau, \xi))$ is a solution of

$$\dot{x} = A(t)\, x + G(t, x + \varphi(t; \tau, \xi)) - F(t, \varphi(t; \tau, \xi)). \tag{13}$$

Observation 1 is a simple but powerful consequence of Lemma 2.1. Next we expose a connection between $D_x^\ell h(t, 0)$ and $D_\xi^\ell [h(t, \varphi(t; \tau, \xi))]|_{\xi=0}$.

Observation 2. For all $t, \tau \in \mathbb{R}$, $\eta = (\eta_1, ..., \eta_n) \in E_1 \times \cdots \times E_n = \mathbb{R}^N$

$$D_\xi^\ell [h(t, \varphi(t; \tau, \xi))]|_{\xi=0} \cdot \eta^\ell = D_x^\ell h(t, 0) \cdot [\Phi_1(t, \tau)\, \eta_1]^{\ell_1} \cdots [\Phi_n(t, \tau)\, \eta_n]^{\ell_n}.$$

This can be seen by calculating the partial derivatives which is easily possible since, by (11), the Taylor coefficients of F and G are zero up to order $|\ell|-1$. The evolution operators Φ_i come in play because of

$$D_{\xi_i}\varphi(t;\tau,\xi)|_{\xi=0}=(0,\ldots,0,\Phi_i(t,\tau),0,\ldots,0)\in L(E_i;\mathbb{R}^N).$$

Now we replace the η_i in Observation 2 by $\Phi_i(\tau,t)\,\zeta_i$ and with the identity $[\Phi_i(t,\tau)]^{-1}=\Phi_i(\tau,t)$ we get

Observation 3. For all $t,\tau\in\mathbb{R}$ and $\zeta\in\mathbb{R}^N$ we have

$$D_x^\ell h(t,0)\cdot\zeta^\ell=D_\xi^\ell[h(t,\varphi(t;\tau,\xi))]|_{\xi=0}\cdot[\Phi_1(\tau,t)\,\zeta_1]^{\ell_1}\cdots[\Phi_n(\tau,t)\,\zeta_n]^{\ell_n}.$$

Now we have a relationship between the Taylor coefficient $D_x^\ell h(t,0)$ of h and the partial derivative $D_\xi^\ell[h(t,\varphi(t;\tau,\xi))]|_{\xi=0}$. By Observation 1 $h(t,\varphi(t;\tau,\xi))$ is a solution of (13). Then by differentiation one can show

Observation 4. The function $D_\xi^\ell[h(t,\varphi(t;\tau,\xi))]_{\xi=0}$ is a solution of

$$\dot{x}=A(t)\,x+c(t), \tag{14}$$

the variational equation of (13) in $L^\ell(E_1,\ldots,E_n;\mathbb{R}^N)$, where

$$c(t)=-(0,\ldots,0,D_x^\ell F_j(t,0),0,\ldots,0)\cdot[\Phi_1(t,\tau)]^{\ell_1}\cdots[\Phi_n(t,\tau)]^{\ell_n}.$$

So far we have assumed that a C^k equivalence $H(t,x)=x+h(t,x)$ between (10) and (12) exists and by Observation 3 the Taylor coefficient $\frac{1}{\ell!}D_x^\ell h(t,0)\cdot x^\ell$ is a function of a special solution $D_\xi^\ell[h(t,\varphi(t;\tau,\xi))]|_{\xi=0}$ of the differential equation (14) and the known evolution operators Φ_i.

From now on we want to use this information to construct a candidate for a C^k equivalence between (10) and (12). We make the ansatz

$$H(t,x)=x+\frac{1}{\ell!}D_x^\ell h(t,0)\cdot x^\ell,$$

i.e., $h(t,x)=\frac{1}{\ell!}D_x^\ell h(t,0)\cdot x^\ell$ has only one nontrivial Taylor coefficient.

We make use of Observations 3 and 4 in the way that we choose a special solution μ of (14) and interpret μ as $D_\xi^\ell[h(t,\varphi(t;\tau,\xi))]|_{\xi=0}$. With Observation 3 and our ansatz this yields

$$H(t,x)=x+\frac{1}{\ell!}\mu(t)\cdot[\Phi_1(\tau,t)\,x_1]^{\ell_1}\cdots[\Phi_n(\tau,t)\,x_n]^{\ell_n}. \tag{15}$$

Which solution μ of (14) should we choose? To satisfy the condition (C) of Definition 2.2 it is necessary that $\lim_{x\to 0}H(t,x)=0$ uniformly in $t\in\mathbb{R}$ and this is satisfied if $\mu(t)\cdot[\Phi_1(\tau,t)\cdot]^{\ell_1}\cdots[\Phi_n(\tau,t)\cdot]^{\ell_n}$ is bounded for $t\in\mathbb{R}$.

Using Siegmund [52] one knows the exponential growth rates of the evolution operators Φ_i of $\dot{x}_i = A_i(t)\, x_i$. Now it is the exponential growth rate of μ we have to take care of. Here a key lemma comes in play.

LEMMA 3.2. *Consider the jth component of* (14)

$$\dot{x}_j = A_j(t)\, x_j + c_j(t). \tag{16}$$

(A) *Assume the condition*

$$a_j > \ell_1 b_1 + \cdots + \ell_n b_n\,. \tag{17}$$

Choose a $\gamma \in (\ell_1 b_1 + \cdots + \ell_n b_n, a_j)$. *Then* $\mu_j(t) := -\int_t^\infty \Phi_j(t, s)\, c_j(s)\, ds$ *is the unique solution of* (16) *with the exponential growth rate* γ *for* $t \to \infty$, *i.e.* $\|\mu_j(t)\| \leqslant C' e^{\gamma t}$ *for all* $t \geqslant 0$ *with some* $C' \geqslant 0$.

(B) *Assume the condition*

$$b_j < \ell_1 a_1 + \cdots + \ell_n a_n\,. \tag{18}$$

Choose a $\gamma \in (b_j, \ell_1 a_1 + \cdots + \ell_n a_n)$. *Then* $\mu_j(t) := \int_{-\infty}^t \Phi_j(t, s)\, c_j(s)\, ds$ *is the unique solution of* (16) *with the exponential growth rate* γ *for* $t \to -\infty$.

Proof. Use Siegmund [52] to show that for $\varepsilon > 0$ exists a $K \geqslant 1$ with

$$\begin{aligned} \|c_j(t)\| &\leqslant \|D_x^\ell F_j(t, 0)\| \cdot \|\Phi_1(t, \tau)\|^{\ell_1} \cdots \|\Phi_n(t, \tau)\|^{\ell_n} \\ &\leqslant MK^{|\ell|} \begin{cases} e^{(\ell_1 b_1 + \cdots + \ell_n b_n + \varepsilon)(t-\tau)} & \text{for a.a.} \quad t \geqslant \tau \\ e^{(\ell_1 a_1 + \cdots + \ell_n a_n - \varepsilon)(t-\tau)} & \text{for a.a.} \quad t \leqslant \tau \end{cases}. \end{aligned}$$

The rest follows from Aulbach and Wanner [6, Lemma 3.2, p. 63]. ∎

Now we assume that (17) or (18) holds (both together can not hold) and choose the following special solution $\mu = (\mu_1, ..., \mu_n)$ of (14)

$$\mu_i(t) := \begin{cases} 0 & \text{if} \quad i \neq j\,, \\ -\int_t^\infty \Phi_j(t, s)\, c_j(s)\, ds & \text{if} \quad i = j \text{ and (17) holds}\,, \\ \int_{-\infty}^t \Phi_j(t, s)\, c_j(s)\, ds & \text{if} \quad i = j \text{ and (18) holds}\,. \end{cases}$$

Using (15) and the identity $\Phi_j(s,\tau)\cdot\Phi_j(\tau,t)=\Phi_j(s,t)$ our explicit candidate $H(t,x)=x+h(t,x)$ for a C^k equivalence is defined by

$$h_i(t,x)=\begin{cases} 0 & \text{if } \ i\neq j, \\ \int_t^\infty \Phi_j(t,s)\,\frac{1}{\ell!}\,D_x^\ell F_j(s,0)\cdot[\Phi_1(s,t)\,x_1]^{\ell_1}\cdots[\Phi_n(s,t)\,x_n]^{\ell_n}\,ds, & \\ \qquad \text{if } \ i=j \text{ and (17) holds}, & \\ -\int_{-\infty}^t \Phi_j(t,s)\,\frac{1}{\ell!}\,D_x^\ell F_j(s,0)\cdot[\Phi_1(s,t)\,x_1]^{\ell_1}\cdots[\Phi_n(s,t)\,x_n]^{\ell_n}\,ds, & \\ \qquad \text{if } \ i=j \text{ and (18) holds.} & \end{cases} \tag{19}$$

Let us have a closer look at the two conditions (17) and (18). For two compact intervals $[a,b]$ and $[c,d]$ we introduce an addition $[a,b]+[c,d]:=[a+b,c+d]$ and for $\gamma\in\mathbb{R}$ a multiplication $\gamma\cdot[a,b]:=[\gamma a,\gamma b]$, furthermore we will use the relation $[a,b]<[c,d]:\Leftrightarrow b<c$ and analog for $[a,b]>[c,d]$. With this notation the conditions (17) and (18) are equivalent to

$$\lambda_j>\sum_{i=1}^n \ell_i\lambda_i \qquad \text{resp.} \qquad \lambda_j<\sum_{i=1}^n \ell_i\lambda_i\,.$$

So for our explicit candidate of H to be well-defined, we have to assume that one of these two conditions is satisfied and this is equivalent to the so-called *nonresonance condition*

$$\boxed{\lambda_j\cap\sum_{i=1}^n \ell_i\lambda_i=\varnothing\,.} \tag{20}$$

If the condition (20) does not hold, then we have a *resonance of order* $|\ell|$ and the term $(0,\ldots,0,\frac{1}{\ell!}\,D_x^\ell F_j(t,0)\cdot x^\ell,0,\ldots,0)$ is called *resonant*.

If the linear part A of the system (10) does not depend on t, then the dichotomy spectrum $\Sigma(A)$ consists of the real parts $\lambda_1,\ldots,\lambda_n$ of the eigenvalues of A and the nonresonance condition (20) for the spectral intervals reduces to Poincarés nonresonance condition (1) for real eigenvalues.

Before we proof that H indeed is a C^k equivalence we want to understand it by the example from the beginning. Therefore consider again

$$\boxed{\begin{aligned}\dot{x}&=x\\ \dot{y}&=\lambda y+x^2\end{aligned}}\,.$$

The spectral intervals of the first and second equation are the one-point sets $\lambda_1=\{1\}$ resp. $\lambda_2=\{\lambda\}$ consisting of the eigenvalues of the linear part.

We want to eliminate the quadratic term x^2 in the second component of the differential equation, i.e. $j=2$ and $\ell=(2,0)$. For $\lambda>2$ the condition $\lambda_2>(2\lambda_1+0\lambda_2)$ holds, so we have no resonance and get

$$\begin{aligned} h_2(t,x,y) &= \int_t^\infty \Phi_2(t,s)\cdot[\Phi_1(s,t)\,x]^2\,ds \\ &= \int_t^\infty e^{\lambda(t-s)}e^{2(s-t)}x^2\,ds = \frac{1}{\lambda-2}\,x^2 \end{aligned}$$

and therefore H is (we get the same h_2 for $\lambda<2$)

$$H(t,x,y)=\begin{pmatrix} x \\ y\end{pmatrix}+\begin{pmatrix} 0 \\ h_2(t,x,y)\end{pmatrix}=\begin{pmatrix} x \\ y+\dfrac{1}{\lambda-2}\,x^2\end{pmatrix}.$$

This is the same result as we calculated above with Poincarés method.

Now we prove that H is indeed a C^k equivalence which eliminates a nonresonant term. In contrast to the condition (11) in the motivation we now allow the right-hand side of the differential equation to have nontrivial Taylor coefficients of arbitrary order.

THEOREM 3.1. *Consider the differential equation* (10). *Let $j\in\{1,\dots,n\}$ be an index and $\ell\in\mathbb{N}_0^n$, $2\leqslant|\ell|\leqslant k$, a multi index. Assume that the nonresonance condition* (20) *holds. Then a local C^k equivalence H exists which eliminates the jth Taylor component $\frac{1}{\ell!}D_x^\ell F_j(t,0)\cdot x^\ell$ belonging to the multi index ℓ and leaves fixed all other Taylor coefficients up to order $|\ell|$.*

That is, (10) *is locally C^k equivalent to a differential equation*

$$\boxed{\dot{x}=A(t)\,x+G(t,x),} \tag{21}$$

with zero reference solution, where G is defined on the set $\mathbb{R}\times B_q(0)$ with $q=\frac{1}{12}\min\left\{p,\frac{\operatorname{Dist}(\lambda_j,\sum_{i=1}^n\ell_i\lambda_i)}{16n^2MK^{|\ell|+1}(p/2)^{|\ell|-2}}\right\}$ and $K=K(j,\ell,A)\geqslant 1$ and for all $\kappa\in\mathbb{N}_0^n$ with $1\leqslant|\kappa|\leqslant|\ell|$ and all $i\in\{1,\dots,n\}$ the identity

$$D_x^\kappa G_i(t,0)\equiv\begin{cases} D_x^\kappa F_i(t,0), & \text{for}\quad \kappa\neq\ell \quad\text{or}\quad i\neq j\\ 0, & \text{for}\quad \kappa=\ell\quad\text{and}\quad i=j\end{cases}$$

holds. The local near-identity C^k equivalence $H\colon\mathbb{R}\times B_{p'}(0)\to B_q$, $(t,x)\mapsto x+h(t,x)$, $p':=\frac{3}{4}q$, between (10) *and* (21) *with respect to the zero solutions* *is defined through* (19). *The inverse transformation* $H^{-1}\colon \mathbb{R}\times B_{q'}(0)\to B_p(0)$, $q' := q$, *has the form*

$$H^{-1}(t,x) = x - h(t,x) + \psi(t,x)$$

with a continuous mapping $\psi\colon \mathbb{R}\times B_{q'}(0)\to\mathbb{R}^N$ *which satisfies the limiting relation* $\lim_{x\to 0}\frac{\psi(t,x)}{\|x\|^{|\ell|^2-1}} = 0$ *uniformly in* $t\in\mathbb{R}$. *Moreover for every* $t\in\mathbb{R}$ *one has the estimates*

$$\|H(t,x)-H(t,\bar{x})\| \leqslant \frac{2^{|\ell|+3}\,n+1}{2^{|\ell|+3}\,n}\,\|x-\bar{x}\| \qquad \textit{for all } x, \bar{x}\in B_{p'}(0),$$

$$\|H^{-1}(t,x)-H^{-1}(t,\bar{x})\| \leqslant \frac{2^{|\ell|+3}\,n}{2^{|\ell|+3}\,n-1}\,\|x-\bar{x}\| \qquad \textit{for all } x, \bar{x}\in B_{q'}(0).$$

Remark 3.1. Note that Theorem 3.1 remains valid, if two blocks A_i and A_j of the linearization $\dot{x} = \operatorname{diag}(A_1(t), ..., A_n(t))\,x$ have the same dichotomy spectrum $\Sigma(A_i) = \Sigma(A_j)$. We will use this fact in the example of the *Duffing–van der Pol oscillator* under nonautonomous parametric perturbation at the end of this article.

Proof of Theorem 3.1. The proof is divided into 12 steps. In the first step we show that h is well-defined. The smoothness of H is examined in the second and third step. In the fourth step we construct the inverse transformation H^{-1} and in the following step the Lipschitz estimates for H and H^{-1} are shown. The smoothness and explicit form of H^{-1} is elaborated in steps 6 to 8. The differential equation $\dot{x} = A(t)\,x + G(t,x)$ is constructed in the ninth step and the following two steps are used to show that G is a C^k Carathéodory function and coincides up to order $|\ell|$ with F except for the jth component of the Taylor component belonging to the multi index ℓ, this component is eliminated in G. In the final step it is proved that H is a local C^k equivalence between (10) and (21) with respect to the zero reference solutions.

STEP 1. *The mapping* $h\colon \mathbb{R}\times\mathbb{R}^N\to\mathbb{R}^N$ *is well-defined and the estimate*

$$\|h(t,x)\| \leqslant \frac{2MK^{|\ell|+1}}{\ell!\,\operatorname{Dist}(\lambda_j, \sum_{i=1}^n \ell_i\lambda_i)}\,\|x_1\|^{\ell_1}\cdots\|x_n\|^{\ell_n} \tag{22}$$

holds with a constant $K = K(j,\ell,A)\geqslant 1$.

Proof of Step 1: The nonresonance condition (20) implies one of the two estimates (17) or (18) for the spectral intervals $\lambda_i = [a_i, b_i]$, $i = 1, ..., n$.

In case of (17) one has $\mathrm{Dist}(\lambda_j, \sum_{i=1}^n \ell_i\lambda_i) = a_j - (\ell_1 b_1 + \cdots + \ell_n b_n)$ and in case (18) $\mathrm{Dist}(\lambda_j, \sum_{i=1}^n \ell_i\lambda_i) = \ell_1 a_1 + \cdots + \ell_n a_n - b_j$ holds. We define $\varepsilon := \frac{\mathrm{Dist}(\lambda_j, \sum_{i=1}^n \ell_i\lambda_i)}{2(|\ell|+1)}$ and choose for every spectral interval $\lambda_i = [a_i, b_i]$ two numbers α_i and β_i with

$$a_i - \varepsilon \leqslant \alpha_i < a_i \qquad \text{and} \qquad b_i < \beta_i \leqslant b_i + \varepsilon.$$

Then as a consequence of Siegmund [52] a $K = K(j, \ell, A) \geqslant 1$ exists with

$$\begin{aligned} \|\Phi_i(t, s)\| &\leqslant K e^{\beta_i(t-s)} && \text{for} \quad t \geqslant s \\ \|\Phi_i(t, s)\| &\leqslant K e^{\alpha_i(t-s)} && \text{for} \quad t \leqslant s \end{aligned}$$

for $i = 1, ..., n$. For all $t \in \mathbb{R}$, $x \in \mathbb{R}^N$ and a.a. $s \in \mathbb{R}$ one has

$$\begin{aligned} &\left\| \Phi_j(t, s) \frac{1}{\ell!} D_x^\ell F_j(s, 0) \cdot [\Phi_1(s, t)\, x_1]^{\ell_1} \cdots [\Phi_n(s, t)\, x_n]^{\ell_n} \right\| \\ &\quad \leqslant \|\Phi_j(t, s)\| \cdot \frac{1}{\ell!} M \cdot \|\Phi_1(s, t)\|^{\ell_1} \cdots \|\Phi_n(s, t)\|^{\ell_n} \cdot \|x_1\|^{\ell_1} \cdots \|x_n\|^{\ell_n} \\ &\quad \leqslant \frac{1}{\ell!} M K^{|\ell|+1} \|x_1\|^{\ell_1} \cdots \|x_n\|^{\ell_n} \begin{cases} e^{\beta_j - (\ell_1\alpha_1 + \cdots + \ell_n\alpha_n)(t-s)}, & \text{if} \quad t \geqslant s \\ e^{\alpha_j - (\ell_1\beta_1 + \cdots + \ell_n\beta_n)(t-s)}, & \text{if} \quad t \leqslant s \end{cases} \end{aligned}$$

In case of (17) the inequality $\alpha_j - (\ell_1\beta_1 + \cdots + \ell_n\beta_n) \geqslant a_j - (\ell_1 b_1 + \cdots + \ell_n b_n) - (|\ell|+1)\, \varepsilon = \frac{1}{2} \mathrm{Dist}(\lambda_j, \sum_{i=1}^n \ell_i\lambda_i)$ yields the estimate

$$\begin{aligned} &\left\| \int_t^\infty \Phi_j(t, s) \frac{1}{\ell!} D_x^\ell F_j(s, 0) \cdot [\Phi_1(s, t)\, x_1]^{\ell_1} \cdots [\Phi_n(s, t)\, x_n]^{\ell_n}\, ds \right\| \\ &\quad \leqslant \frac{1}{\ell!} M K^{|\ell|+1} \|x_1\|^{\ell_1} \cdots \|x_n\|^{\ell_n} \int_t^\infty e^{\frac{1}{2}\mathrm{Dist}(\lambda_j, \sum_{i=1}^n \ell_i\lambda_i)(t-s)}\, ds \\ &\quad = \frac{2MK^{|\ell|+1}}{\ell!\, \mathrm{Dist}\, \lambda_j, \sum_{i=1}^n \ell_i\lambda_i)} \|x_1\|^{\ell_1} \cdots \|x_n\|^{\ell_n} \end{aligned}$$

and (22) follows. Analogously if (18) holds then the inequality $\beta_j - (\ell_1\alpha_1 + \cdots + \ell_n\alpha_n) \leqslant b_j - (\ell_1 a_1 + \cdots + \ell_n a_n) + (|\ell|+1)\, \varepsilon = -\frac{1}{2} \mathrm{Dist}(\lambda_j, \sum_{i=1}^n \ell_i\lambda_i)$ implies the estimate (22).

STEP 2. *The mapping $H: \mathbb{R} \times \mathbb{R}^N \to \mathbb{R}^N$, $(t, x) \mapsto x + h(t, x)$, is continuous and infinitely many times continuously partially differentiable in x.*

Proof of Step 2. Arguing for each component seperately, the proof of this claim reduces to the verification that $h_j\colon \mathbb{R}\times\mathbb{R}^N \to E_j$ is continuous and infinitely many times continuously partially differentiable in x. An application of Lebesgues dominated convergence theorem yields this continuity and a similar argument proves the differentiability of h_j in x and one gets for all $t \in \mathbb{R}$ and $x, \xi \in \mathbb{R}^N$ the identity

$$\begin{aligned}
&D_x h_j(t, x)\cdot\xi \\
&\quad = \int_t^\infty D_x\left[\Phi_j(t, s)\frac{1}{\ell!}D_x^\ell F_j(s, 0)\cdot[\Phi_1(s, t)\, x_1]^{\ell_1}\cdots[\Phi_n(s, t)\, x_n]^{\ell_n}\right]\cdot\xi\, ds \\
&\quad = \sum_{i=1, ..., n:\, \ell_i \geqslant 1} \ell_i \int_t^\infty \Phi_j(t, s)\frac{1}{\ell!}D_x^\ell F_j(s, 0) \\
&\qquad \cdot[\Phi_1(s, t)\, x_1]^{\ell_1}\cdots[\Phi_i(s, t)\,\xi_i]\cdot[\Phi_i(s, t)\, x_i]^{\ell_i-1}\cdots[\Phi_n(s, t)\, x_n]^{\ell_n}\, ds.
\end{aligned} \tag{23}$$

The continuity and differentiability of (24) in x follows again from Lebesgues theorem. The second derivative operates on $\xi, \eta \in \mathbb{R}^N$ through

$$\begin{aligned}
&D_x^2 h_j(t, x)\cdot\xi\cdot\eta \\
&\quad = \sum_{i, m=1, ..., n:\, \ell_i, \ell_m \geqslant 1, i\neq m} \ell_i\ell_m \int_t^\infty \Phi_j(t, s)\frac{1}{\ell!}D_x^\ell F_j(s, 0) \\
&\qquad \cdot[\Phi_1(s, t)\, x_1]^{\ell_1}\cdots[\Phi_i(s, t)\,\xi_i]\cdot[\Phi_i(s, t)\, x_i]^{\ell_i-1} \\
&\qquad \cdots[\Phi_m(s, t)\,\eta_m]\cdot[\Phi_m(s, t)\, x_m]^{\ell_m-1}\cdots[\Phi_n(s, t)\, x_n]^{\ell_n}\, ds \\
&\qquad + \sum_{i=1, ..., n:\, \ell_i \geqslant 2} \ell_i(\ell_i-1)\int_t^\infty \Phi_j(t, s)\frac{1}{\ell!}D_x^\ell F_j(s, 0)\cdot[\Phi_1(s, t)\, x_1]^{\ell_1} \\
&\qquad \cdots[\Phi_i(s, t)\,\xi_i]\cdot[\Phi_i(s, t)\,\eta_i]\cdot[\Phi_i(s, t)\, x_i]^{\ell_i-2}\cdots[\Phi_n(s, t)\, x_n]^{\ell_n}\, ds.
\end{aligned} \tag{24}$$

Mathematical induction yields the existence and continuity of the partial derivatives $D_x^m h_j\colon \mathbb{R}\times\mathbb{R}^N \to L^m(\mathbb{R}^N; E_j)$ for $m = 1, ..., k$. For $m > k$ the mapping $D_x^m h_j$ is zero and for this reason h_j and therefore also H is continuous and infinitely many times continuously partially differentiable with respect to x.

STEP 3. *The partial derivative $D_t H(t, x)$ exists for a.a. $t \in \mathbb{R}$ and for all $x \in \mathbb{R}^N$ and $D_t H\colon \mathbb{R}\times\mathbb{R}^N \to \mathbb{R}^N$ is a C^k Carathéodory function.*

Proof of Step 3. We prove the claim in case of (17). The integral $h_j(t,x) = \int_t^\infty \Phi_j(t,s) \frac{1}{\ell!} D_x F_j(t,0) \cdot [\Phi_1(s,t)\, x_1]^{\ell_1} \cdots [\Phi_n(s,t)\, x_n]^{\ell_n}\, ds$ is partially differentiable in t for a.a. $t \in \mathbb{R}$ and for all $x \in \mathbb{R}^N$ with the derivative

$$D_t h_j(t,x) = -\Phi_j(t,t) \frac{1}{\ell!} D_x^\ell F_j(t,0) \cdot [\Phi_1(t,t)\, x_1]^{\ell_1} \cdots [\Phi_n(t,t)\, x_n]^{\ell_n}$$

$$+ \int_t^\infty D_t \left[\Phi_j(t,s) \frac{1}{\ell!} D_x^\ell F_j(t,0) \cdot [\Phi_1(s,t)\, x_1]^{\ell_1} \right.$$

$$\left. \cdots [\Phi_n(s,t)\, x_n]^{\ell_n} \right] ds.$$

Using the relation $D_t \Phi(s,t) = -\Phi(s,t)\, A(t)$ one has

$$D_t h_j(t,x) = -\frac{1}{\ell!} D_x^\ell F_j(t,0) \cdot x^\ell + A_j(t)\, h_j(t,x)$$

$$+ \sum_{i=1,\ldots,n:\, \ell_i \geqslant 1} \ell_i \int_t^\infty \Phi_j(t,s) \frac{1}{\ell!} D_x^\ell F_j(t,0)$$

$$\cdot [\Phi_1(s,t)\, x_1]^{\ell_1} \cdots [-\Phi_i(s,t)\, A_i(t)\, x_i] \cdot [\Phi_i(s,t)\, x_i]^{\ell_i - 1}$$

$$\cdots [\Phi_n(s,t)\, x_n]^{\ell_n}\, ds. \tag{25}$$

Aulbach and Wanner [6, Lemma 2.2, p. 49] implies that the first two summands are C^k Carathéodory functions and similarly for the third.

Step 4. *Define* $q := \frac{1}{12} \min\{p, \frac{\operatorname{Dist}(\lambda_j, \sum_{i=1}^n \ell_i \lambda_i)}{16 n^2 M K^{|\ell|+1} (p/2)^{|\ell|-2}}\}$, $p' := \frac{3}{4} q$ *and* $q' := q$. *Choose and fix an arbitrary* $t \in \mathbb{R}$. *Then the restriction*

$$H(t,\cdot) \colon B_{4p'}(0) \subset B_p(0) \to H(t, B_{4p'}(0))$$

of $H(t,\cdot)$ *is a* C^k *diffeomorphism and for every* $t \in \mathbb{R}$ *a* C^k *diffeomorphism*

$$H^{-1}(t,\cdot) \colon B_{q'}(0) \to H^{-1}(t, B_{q'}(0))$$

exists such that for all $x \in B_{4p'}(0)$ *with* $H(t,x) \in B_{q'}(0)$ *the identity*

$$H^{-1}(t, H(t,x)) = x$$

holds. Moreover for all $x \in B_{q'}(0)$ *we have* $H^{-1}(t,x) \in B_{4p'}(0)$ *and*

$$H(t, H^{-1}(t,x)) = x.$$

Proof of Step 4. Formula (24) implies for all $t \in \mathbb{R}$ and $x \in B_{\frac{\rho}{2}}(0)$

$$\begin{aligned}
\|D_x^2 H(t,x)\| \leqslant & \sum_{i,m=1,\ldots,n:\, \ell_i,\ell_m \geqslant 1,\, i \neq m} \ell_i \ell_m \frac{2MK^{|\ell|+1}}{\ell!\,\mathrm{Dist}(\lambda_j, \sum_{i=1}^n \ell_i\lambda_i)} \\
& \cdot \|x_1\|^{\ell_1} \cdots \|x_i\|^{\ell_i-1} \cdots \|x_m\|^{\ell_m-1} \cdots \|x_n\|^{\ell_n} \\
& + \sum_{i=1,\ldots,n:\, \ell_i \geqslant 2} \ell_i(\ell_i-1) \frac{2MK^{|\ell|+1}}{\ell!\,\mathrm{Dist}(\lambda_j, \sum_{i=1}^n \ell_i\lambda_i)} \\
& \cdot \|x_1\|^{\ell_1} \cdots \|x_i\|^{\ell_i-2} \cdots \|x_n\|^{\ell_n} \leqslant \\
\leqslant & \frac{2n^2MK^{|\ell|+1}}{\mathrm{Dist}(\lambda_j, \sum_{i=1}^n \ell_i\lambda_i)} \left(\frac{p}{2}\right)^{|\ell|-2}.
\end{aligned}$$

In case of (18) one can show analogously the same estimate. Now Abraham, Marsden, Ratiu [1, Proposition 2.5.6, pp. 119–121] implies the claim.

STEP 5. *For every* $t \in \mathbb{R}$ *we have*

$$\|H(t,x) - H(t,\bar{x})\| \leqslant \frac{2^{|\ell|+3}\, n+1}{2^{|\ell|+3}\, n} \|x - \bar{x}\| \qquad \textit{for} \quad x, \bar{x} \in B_{4p'}(0), \quad (26)$$

$$\|H^{-1}(t,x) - H^{-1}(t,\bar{x})\| \leqslant \frac{2^{|\ell|+3}\, n}{2^{|\ell|+3}\, n-1} \|x - \bar{x}\| \qquad \textit{for} \quad x, \bar{x} \in B_{q'}(0). \quad (27)$$

Proof of Step 5. First we prove the Lipschitz continuity of $h\colon \mathbb{R} \times B_{4p'}(0) \to \mathbb{R}^N$ in the second argument. The formula (23) for the partial derivative of h_j w.r.t. x implies for all $t \in \mathbb{R}$ and $x \in B_{4p'}(0)$ the estimate

$$\begin{aligned}
&\|D_x h(t,x)\| \\
&\leqslant \sum_{i=1,\ldots,n:\, \ell_i \geqslant 1} \ell_i \frac{2MK^{|\ell|+1}}{\ell!\,\mathrm{Dist}(\lambda_j, \sum_{i=1}^n \ell_i\lambda_i)} \|x_1\|^{\ell_1} \cdots \|x_i\|^{\ell_i-1} \cdots \|x_n\|^{\ell_n} \\
&\leqslant n \frac{2MK^{|\ell|+1}}{\mathrm{Dist}(\lambda_j, \sum_{i=1}^n \ell_i\lambda_i)} \cdot (4p') \cdot (4p')^{|\ell|-2} \\
&\leqslant \frac{1}{32n\left(\frac{p}{2}\right)^{|\ell|-2}} \left(\frac{p}{4}\right)^{|\ell|-2} \leqslant \frac{1}{2^{|\ell|+3}\, n} < 1.
\end{aligned}$$

This yields for all $t \in \mathbb{R}$ and $x, \bar{x} \in B_{4p'}(0)$ the estimate

$$\|h(t,x) - h(t,\bar{x})\| \leqslant \frac{1}{2^{|\ell|+3}\, n} \|x - \bar{x}\| \qquad (28)$$

and this implies (26). To prove the Lipschitz estimate for H^{-1} we use (28) to show for $t \in \mathbb{R}$ and $y, \bar{y} \in B_{4p'}(0)$ the estimate

$$\|y-\bar{y}\| - \frac{1}{2^{|\ell|+3}\,n}\|y-\bar{y}\| \leqslant \|y-\bar{y}\| - \|h(t,y)-h(t,\bar{y})\|$$

and it follows that

$$\frac{2^{|\ell|+3}\,n-1}{2^{|\ell|+3}\,n}\|y-\bar{y}\| \leqslant \|H(t,y)-H(t,\bar{y})\|.$$

Step 4 implies for $x, \bar{x} \in B_{q'}(0)$ the inclusions $y := H^{-1}(t,x) \in B_{4p'}(0)$, $\bar{y} := H^{-1}(t,\bar{x}) \in B_{4p'}(0)$ and one gets the estimate (27).

STEP 6. *The mapping* $H^{-1}\colon \mathbb{R} \times B_{q'}(0) \to \mathbb{R}^N$ *is continuous.*

Proof of Step 6. Let $t_0 \in \mathbb{R}$ and $x_0 \in B_{q'}(0)$ be arbitrary, but fixed. For $t \in \mathbb{R}$ and $x \in B_{q'}(0)$ formula (27) implies the estimate

$$\begin{aligned}
&\|H^{-1}(t,x)-H^{-1}(t_0,x_0)\| \\
&\quad \leqslant \|H^{-1}(t,x)-H^{-1}(t,x_0)\| + \|H^{-1}(t,x_0)-H^{-1}(t_0,x_0)\| \\
&\quad \leqslant \frac{2^{|\ell|+3}\,n}{2^{|\ell|+3}\,n-1}\|x-x_0\| + \|H^{-1}(t,x_0)-H^{-1}(t_0,x_0)\|
\end{aligned}$$

and it suffices to show the limiting relation $\lim_{t\to t_0} H^{-1}(t,x_0) = H^{-1}(t_0,x_0)$. Therefore consider the identity

$$x_0 = H(t, H^{-1}(t,x_0)) = H^{-1}(t,x_0) + h(t, H^{-1}(t,x_0))$$

for all $t \in \mathbb{R}$. With (28) we have the estimate

$$\begin{aligned}
&\|H^{-1}(t,x_0)-H^{-1}(t_0,x_0)\| \\
&\quad = \|h(t,H^{-1}(t,x_0)) - h(t_0,H^{-1}(t_0,x_0))\| \\
&\quad \leqslant \|h(t,H^{-1}(t,x_0)) - h(t,H^{-1}(t_0,x_0))\| \\
&\qquad + \|h(t,H^{-1}(t_0,x_0)) - h(t_0,H^{-1}(t_0,x_0))\| \\
&\quad \leqslant \frac{1}{2^{|\ell|+3}\,n}\|H^{-1}(t,x_0)-H^{-1}(t_0,x_0)\| \\
&\qquad + \|h(t,H^{-1}(t_0,x_0)) - h(t_0,H^{-1}(t_0,x_0))\|.
\end{aligned}$$

Solving for $\|H^{-1}(t, x_0) - H^{-1}(t_0, x_0)\|$ yields

$$\|H^{-1}(t, x_0) - H^{-1}(t_0, x_0)\| \leqslant \left[1 - \frac{1}{2^{|\ell|+3}\, n}\right]^{-1} \|h(t, H^{-1}(t_0, x_0)) - h(t_0, H^{-1}(t_0, x_0))\|$$

and with the continuity of h (step 2) the claim follows.

STEP 7. *The mapping* $H^{-1}\colon \mathbb{R} \times B_{q'}(0) \to \mathbb{R}^N$ *is* k*-times continuously partially differentiable w.r.t.* x.

Proof of Step 7. Due to step 4, the partial derivative

$$D_x^m H^{-1}\colon \mathbb{R} \times B_{q'}(0) \to L^m(\mathbb{R}^N)$$

exists for every $m \in \{1, ..., k\}$. We have for every $t \in \mathbb{R}$ and $x \in B_{q'}(0)$

$$D_x H^{-1}(t, x) = [D_x H(t, H^{-1}(t, x))]^{-1},$$

i.e., the mapping $D_x H^{-1}\colon \mathbb{R} \times B_{q'}(0) \to \mathcal{GL}(\mathbb{R}^N)$ has a representation as the composition

$$\mathbb{R} \times B_{q'}(0) \xrightarrow{\mathrm{proj}_1 \times H^{-1}} \mathbb{R} \times B_{4p'}(0) \xrightarrow{D_x H} \mathcal{GL}(\mathbb{R}^N) \xrightarrow{[\,\cdot\,]^{-1}} \mathcal{GL}(\mathbb{R}^N).$$

The function H^{-1} is continuous (step 6) and also $D_x H$ is continuous (step 2). Due to Abraham, Marsden, Ratiu [1, Lemma 2.2.5, p. 117] the inverse $[\,\cdot\,]^{-1}\colon \mathcal{GL}(\mathbb{R}^N) \to \mathcal{GL}(\mathbb{R}^N)$ is C^∞. Therefore the mapping $D_x H^{-1}$ is, as a composition of continuous mappings, continuous itself. Now let $m \in \{2, ..., k\}$ and let the continuity of $D_x^i H^{-1}$ be proved for $1 \leqslant i \leqslant m-1$. It is easy to check that the mapping $D_x^m H^{-1}\colon \mathbb{R} \times B_{q'}(0) \to L^m(\mathbb{R}^N)$ has the form

$$D_x^m H^{-1}(t, x) = -[D_x H(t, H^{-1}(t, x))]^{-1} \cdot S(t, x),$$

where $S(t, x)$ is a sum with summands of the form

$$D_x^m H(t, H^{-1}(t, x)) \cdot D_x^{\ell_1} H^{-1}(t, x) \cdots D_x^{\ell_u} H^{-1}(t, x).$$

Here $u = 2, ..., m$ and $\ell \in \mathbb{N}^u$ with $|\ell| = m$, i.e. $\ell_1, ..., \ell_u \leqslant m-1$. The assumption of the induction as well as step 2 and step 6 imply that all mappings are continuous and therefore the mapping $D_x^m H^{-1}$ is also continuous if one recalls that for $i = 1, ..., u$ the evaluation mapping

$$(\mathbb{R}^N)^{\ell_i} \times L^{\ell_i}(\mathbb{R}^N) \to \mathbb{R}^N, \quad (\xi_1, ..., \xi_{\ell_i}, T) \mapsto T \cdot (\xi_1, ..., \xi_{\ell_i})$$

is continuous.

STEP 8. *The mapping* $H^{-1}\colon \mathbb{R}\times B_{q'}(0)\to\mathbb{R}^N$ *is of the form*

$$H^{-1}(t,x)=x-h(t,x)+\psi(t,x)$$

with a continuous mapping $\psi\colon \mathbb{R}\times B_{q'}(0)\to\mathbb{R}^N$ *that satisfies*

$$\lim_{x\to 0}\frac{\|\psi(t,x)\|}{\|x\|^{|\ell|^2-1}}=0 \quad \textit{uniformly in } t\in\mathbb{R}. \tag{29}$$

Proof of Step 8. The inverse of $H(t,\cdot)$ can be given explicitely. Therefore let $t\in\mathbb{R}$ be arbitrary but fixed. For every $x\in B_{4p'}(0)$ one has (with $|\ell|\geqslant 2$, $n\geqslant 1$, $h(t,0)=0$) from the inequality (28) the estimate

$$\|h(t,x)\|\leqslant\tfrac{1}{32}\,\|x\|. \tag{30}$$

We define for $i\in\mathbb{N}_0$ the iteration $(-h)^i(t,\cdot): B_{4p'}(0)\to\mathbb{R}^N$ through

$$\begin{aligned}(-h)^0\,(t,\cdot) &:\equiv \mathrm{id}_{B_{4p'}(0)}\\ (-h)^{i+1}\,(t,\cdot) &:\equiv [-h(t,\cdot)]\circ[(-h)^i(t,\cdot)]\equiv\; -h(t,(-h)^i(t,\cdot)).\end{aligned}$$

Then for all $i\in\mathbb{N}_0$ and $x\in B_{4p'}(0)$ the estimate $\|(-h)^i(t,x)\|\leqslant(\frac{1}{32})^n\,\|x\|$ holds and this implies for every $m\in\mathbb{N}$

$$\left\|\sum_{i=0}^{m}(-h)^i\,(t,x)\right\|\leqslant\sum_{i=0}^{\infty}\|(-h)^i(t,x)\|\leqslant\frac{1}{1-\frac{1}{32}}\,\|x\|=\tfrac{32}{31}\,\|x\|. \tag{31}$$

Therefore for all $x\in B_{4p'}(0)$ the series $\sum_{i=0}^{\infty}(-h)^i(t,x)$ converges absolutely in $\mathbb{R}^N$ and with (31) we have the inclusion

$$\sum_{i=0}^{\infty}(-h)^i\,(t,x)\in B_{4p'}(0)\qquad\text{for all}\quad x\in B_{q'}(0).$$

For $x\in B_{q'}(0)$ also

$$\begin{aligned}&[\mathrm{id}+h(t,\cdot)]\circ\left[\sum_{i=0}^{\infty}(-h)^i\,(t,\cdot)\right](x)\\ &\quad=x=\left[\sum_{i=0}^{\infty}(-h)^i\,(t,\cdot)\right]\circ[\mathrm{id}+h(t,\cdot)](x)\end{aligned}$$

and one has the identity

$$H^{-1}(t,x) := \sum_{i=0}^{\infty} (-h)^i\,(t,x) \qquad \text{for all } t \in \mathbb{R} \qquad \text{and} \qquad x \in B_{q'}(0).$$

With the mapping $\psi\colon \mathbb{R} \times B_{q'}(0) \to \mathbb{R}^N$, $(t,x) \mapsto \sum_{i=2}^{\infty} (-h)^i(t,x)$, one gets for every arbitrary $t \in \mathbb{R}$ and each $x \in B_{q'}(0)$ the identity

$$H^{-1}(t,x) = x - h(t,x) + \psi(t,x).$$

To show the limiting relation (29) one considers the following estimate which is an implication of (22)

$$\|h(t,x)\| \leqslant \frac{2MK^{|\ell|+1}}{\operatorname{Dist}(\lambda_j, \sum_{i=1}^{n} \ell_i \lambda_i)}\, \|x\|^{|\ell|}.$$

Applying the estimate twice together with (30) yields for all $t \in \mathbb{R}$, $x \in B_{q'}(0)$ and $i \geqslant 2$ the inequalities

$$\begin{aligned} \|(-h)^i\,(t,x)\| &\leqslant \frac{2MK^{|\ell|+1}}{\operatorname{Dist}(\lambda_j, \sum_{i=1}^{n} \ell_i \lambda_i)}\, \|(-h)^{i-1}\,(t,x)\|^{|\ell|} \\ &\leqslant \left[\frac{2MK^{|\ell|+1}}{\operatorname{Dist}(\lambda_j, \sum_{i=1}^{n} \ell_i \lambda_i)}\right]^{|\ell|+1} \|(-h)^{i-2}\,(t,x)\|^{|\ell|^2} \\ &\leqslant \left[\frac{2MK^{|\ell|+1}}{\operatorname{Dist}(\lambda_j, \sum_{i=1}^{n} \ell_i \lambda_i)}\right]^{|\ell|+1} \left(\frac{1}{32}\right)^{(i-2)|\ell|^2} \|x\|^{|\ell|^2} \end{aligned}$$

and this implies that

$$\|\psi(t,x)\| \leqslant \left[\frac{2MK^{|\ell|+1}}{\operatorname{Dist}(\lambda_j, \sum_{i=1}^{n} \ell_i \lambda_i)}\right]^{|\ell|+1} \frac{1}{1-\left(\frac{1}{32}\right)^{|\ell|^2}}\, \|x\|^{|\ell|^2}$$

and therefore the limiting relation (29).

STEP 9. *If μ is a solution of* (10)*, which is in $B_{p'}(0)$ then $H(\cdot, \mu(\cdot))$ is a solution of $\dot{x} = \tilde{G}(t,x)$, where the right hand side $\tilde{G}\colon \mathbb{R} \times B_{q'}(0) \to \mathbb{R}^N$ is defined through*

$$D_t H(t, H^{-1}(t,x)) + D_x H(t, H^{-1}(t,x)) \cdot [A(t)\, H^{-1}(t,x) + F(t, H^{-1}(t,x))].$$

If v is a solution of $\dot{x} = \widetilde{G}(t, x)$, which is in $B_{q'}(0)$ then $H^{-1}(\cdot, v(\cdot))$ is a solution of (10).

Proof of Step 9. Let $I \subset \mathbb{R}$ be a nonempty interval and let $\mu: I \to B_{p'}(0)$ be a solution of (10). Now (26) implies (with $|\ell| \geqslant 2$, $n \geqslant 1$, $H(t, 0) = 0$) the estimate $\|H(t, \mu(t))\| \leqslant \frac{33}{32} \|\mu(t)\| \leqslant q' = q$ and therefore the inclusion $H(t, \mu(t)) \in B_q(0)$. For $t, s \in I$ one has

$$\begin{aligned} &H(t, \mu(t)) - H(s, \mu(s)) \\ &\quad = \int_s^t D_\sigma[H(\sigma, \mu(\sigma))]\, ds \\ &\quad = \int_s^t D_t H(\sigma, \mu(\sigma)) + D_x H(\sigma, \mu(\sigma)) \cdot [A(\sigma)\, \mu(\sigma) + F(\sigma, \mu(\sigma))]\, ds \end{aligned}$$

and using the relation $\mu(\sigma) \equiv H^{-1}(\sigma, H(\sigma, \mu(\sigma)))$ we see that $H(\cdot, \mu(\cdot))$: $I \to B_q(0)$ is a solution of the differential equation

$$\begin{aligned} \dot{x} = {} & D_t H(t, H^{-1}(t, x)) \\ & + D_x H(t, H^{-1}(t, x)) \cdot [A(t)\, H^{-1}(t, x) + F(t, H^{-1}(t, x))]. \end{aligned}$$

Now let $v: I \to B_{q'}(0)$ be a solution of $\dot{x} = \widetilde{G}(t, x)$ and let $\tau \in I$ be arbitrary but fixed. As we have seen $H(t, \varphi(t; \tau, H^{-1}(\tau, v(\tau))))$ is the unique solution of the initial value problem $\dot{x} = \widetilde{G}(t, x)$, $x(\tau) = v(\tau)$. Therefore

$$H(t, \varphi(t; \tau, H^{-1}(\tau, v(\tau)))) \equiv v(t) \qquad \text{on} \quad I$$

holds and with step 4 we have the relation $\varphi(t; \tau, H^{-1}(\tau, v(\tau))) \equiv H^{-1}(t, v(t))$ on I. Formula (27) implies the estimate $\|H^{-1}(t, v(t))\| \leqslant \frac{32}{31} \|v(t)\| \leqslant 4p' \leqslant p$ and thus the inclusion $H^{-1}(t, v(t)) \in B_p(0)$. Therefore $H^{-1}(\cdot, v(\cdot)): I \to B_p(0)$ is a solution of (10).

STEP 10. $\widetilde{G}$: $\mathbb{R} \times B_{q'}(0) \to \mathbb{R}^N$ *(see step* 9*) is a C^k Carathéodory function.*

Proof of Step 10. The mapping F is a C^k Carathéodory function. Step 2, 3 and 7 imply for every $t \in \mathbb{R}$ that the mapping $\widetilde{G}(t, \cdot)$: $B_{q'}(0) \to \mathbb{R}^N$ is k-times continuously partially differentiable in x. To show that $\widetilde{G}$ is a C^k Carathéodory function it remains, due to Definition 2.1, to show that for every $j \in \{0, ..., k\}$ and all $x \in B_{q'}(0)$ the mapping $D_x^j \widetilde{G}(\cdot, x)$: $\mathbb{R} \to L^j(\mathbb{R}^N)$

is measurable. For $j=0$ this follows with Aulbach and Wanner [6, Lemma 2.2, p. 49] from Steps 2, 3, and 6. For $j=1$ one calculates for all $t \in \mathbb{R}$ and $x \in B_{\tilde{q}}(0)$ the relation

$$\begin{aligned} D_x \tilde{G}(t,x) = {} & D_x D_t H(t, H^{-1}(t,x)) \cdot D_x H^{-1}(t,x) \\ & + D_x^2 H(t, H^{-1}(t,x)) \cdot D_x H^{-1}(t,x) \\ & \cdot [A(t)\, H^{-1}(t,x) + F(t, H^{-1}(t,x))] \\ & + D_x H(t, H^{-1}(t,x)) \\ & \cdot [A(t)\, D_x H^{-1}(t,x) + D_x F(t, H^{-1}(t,x)) \cdot D_x H^{-1}(t,x)] \end{aligned}$$

and the measurability of $D_x \tilde{G}(\cdot, x)$ follows again with Aulbach and Wanner [6, Lemma 2.2, p. 49] from Steps 2, 3, 6 and 7. Mathematical induction (on the structure of the derivatives $D_x^j \tilde{G}(\cdot, x)$, $j \in \{2, ..., k\}$) proves the claim.

STEP 11. *The mapping $\tilde{G}\colon \mathbb{R} \times B_q \to \mathbb{R}^N$ (see step 9) has the form $\tilde{G}(t,x) = A(t)\, x + G(t,x)$ and for the components of the Taylor coefficients of $G\colon \mathbb{R} \times B_q \to E_1 \times \cdots \times E_n = \mathbb{R}^N$ the following identities hold*

$$D_x^\kappa G_i(t,0) \equiv \begin{cases} D_x^\kappa F_i(t,0), & \text{for } \kappa \neq \ell \quad \text{or} \quad i \neq j \\ 0, & \text{for } \kappa = \ell \quad \text{and} \quad i = j \end{cases}$$

for all $\kappa \in \mathbb{N}_0^n$ with $1 \leqslant |\kappa| \leqslant |\ell|$ and all $i \in \{1, ..., n\}$.

Proof of Step 11. We write the components $\tilde{G}_i\colon \mathbb{R} \times B_q(0) \to E_i$, $i = 1, ..., n$, of the right hand side of the transformed differential equation as a sum of terms up to order $|\ell|$ and terms of higher order. The most important relation to do this is the following connection between $D_t h_j$ and $D_x h_j$. For all $t \in \mathbb{R}$ and $x \in B_q(0)$ the formulas (24) and (26) yield the identity.

$$D_t h_j(t,x) = -\frac{1}{\ell!} D_x^\ell F_j(t,0) \cdot x^\ell + A_j(t)\, h_j(t,x) - D_x h_j(t,x) \cdot [A(t)\, x].$$

The jth component $\tilde{G}_j(t,x)$ of $\tilde{G}$ is

$$D_t H_j(t, H^{-1}(t,x)) + D_x H_j(t, H^{-1}(t,x)) \cdot [A(t)\, H^{-1}(t,x) + F(t, H^{-1}(t,x))]$$

and one gets the identity

$$\begin{aligned}
\tilde{G}_j(t, x) &= -\frac{1}{\ell!} D_x^\ell F_j(t, 0) \cdot [H^{-1}(t, x)]^\ell + A_j(t)\, h_j(t, H^{-1}(t, x)) \\
&\quad - D_x h_j(t, H^{-1}(t, x)) \cdot [A(t)\, H^{-1}(t, x)] \\
&\quad + [\mathrm{proj}_j + D_x h_j(t, H^{-1}(t, x))] \cdot [A(t)\, H^{-1}(t, x) + F(t, H^{-1}(t, x))] \\
&= A_j(t)\, H_j^{-1}(t, x) + F_j(t, H^{-1}(t, x)) - \frac{1}{\ell!} D_x^\ell F_j(t, 0) \cdot [H^{-1}(t, x)]^\ell \\
&\quad + A_j(t)\, h_j(t, H^{-1}(t, x)) + D_x h_j(t, H^{-1}(t, x)) \cdot F(t, H^{-1}(t, x)).
\end{aligned}$$

The explicit form and the estimates of H and H^{-1} yield

$$\tilde{G}_j(t, x) = A_j(t)\, x_j + F_j(t, x) - \frac{1}{\ell!} D_x^\ell F_j(t, 0) \cdot x^\ell + R_j(t, x)$$

with a continuous function $R_j\colon \mathbb{R} \times B_q(0) \to E_j$, which satisfies for every $t \in \mathbb{R}$ the limiting relations $\lim_{x \to 0} \frac{\|R_j(t, x)\|}{\|x\|^{|\ell|}}$. This proves the claim for $\tilde{G}_j$. Now to $\tilde{G}_i$ for $i \neq j$. We have $D_t H_i(t, x) \equiv 0$ and $D_x H_i(t, x) \equiv \mathrm{proj}_i$. With $H_i^{-1}(t, x) = x_i$ this implies the identity

$$\tilde{G}_i(t, x) = A_i(t)\, H_i^{-1}(t, x) + F_i(t, H^{-1}(t, x)) = A_i(t)\, x_i + F_i(t, x) + R_i(t, x)$$

with a continuous function $R_i\colon \mathbb{R} \times B_q(0) \to E_i$ which satisfies for every $t \in \mathbb{R}$ the limiting relation $\lim_{x \to 0} \frac{\|R_i(t, x)\|}{\|x\|^{|\ell|}}$. With the definition $G(t, x) := \tilde{G}(t, x) - A(t)\, x$ the claim follow.

STEP 12. *The mapping* $H\colon \mathbb{R} \times B_{p'}(0) \to B_q(0)$ *is a* C^k *equivalence between the systems* (10) *and* (21) *with respect to the zero solutions with the inverse transformation* $H^{-1}\colon \mathbb{R} \times B_{q'}(0) \to B_p(0)$.

Proof of Step 12. We only have to verify some properties of the definition of a C^k equivalence. Use Steps 4, 9 and 5. ∎

COROLLARY 3.1. *Let A and F be periodic in t with a period $\Theta > 0$, i.e. for all $t \in \mathbb{R}$ and $x \in B_r(0)$ the identities*

$$A(t + \Theta) = A(t) \qquad \text{and} \qquad F(t + \Theta, x) = F(t, x)$$

hold. Then H from Theorem 2 is also periodic in t with period Θ. Especially, if (10) *is autonomous then H is independent of t.*

Proof. For $\tau \in \mathbb{R}$ and $\xi \in \mathbb{R}^N$ the mapping $\Phi(t+\Theta, \tau+\Theta)\,\xi$ is the unique solution of the initial value problem $\dot{x} = A(t+\Theta)\,x$, $x(\tau) = \xi$ and also $\Phi(t, \tau)\,\xi$ is the unique solution of the same initial value problem $\dot{x} = A(t)\,x$, $x(\tau) = \xi$ and therefore the identity $\Phi(t+\Theta, \tau+\Theta) = \Phi(t, \tau)$ holds for all $t, \tau \in \mathbb{R}$. Moreover the Θ periodicity of F in t implies the relation $D_x^\ell F_j(t+\Theta, 0) = D_x^\ell F_j(t, 0)$ and one gets in case of (17) the equality

$$\begin{aligned} h_j(t+\Theta, x) &= \int_{t+\Theta}^{\infty} \Phi_j(t+\Theta, s)\, \frac{1}{\ell!}\, D_x^\ell F_j(s, 0) \\ &\quad \times [\Phi_1(s, t+\Theta)\, x_1]^{\ell_1} \cdots [\Phi_n(s, t+\Theta)\, x_n]^{\ell_n}\, ds \\ &= \int_{t}^{\infty} \Phi_j(t+\Theta, s+\Theta)\, \frac{1}{\ell!}\, D_x^\ell F_j(s+\Theta, 0) \\ &\quad \times [\Phi_1(s+\Theta, t+\Theta)\, x_1]^{\ell_1} \cdots [\Phi_n(s+\Theta, t+\Theta)\, x_n]^{\ell_n}\, ds \\ &= h_j(t, x) \end{aligned}$$

and the claim follows. ∎

Now it is easy to get our main result on normal forms. Combining the three steps we immediately get the following theorem.

Normal Form Theorem *Consider a differential equation*

$$\boxed{\dot{x} = f(t, x)} \tag{32}$$

together with a reference solution $\mu_0 \colon \mathbb{R} \to \mathbb{R}^N$. *Assume that*

(A) $f \colon D \subset \mathbb{R} \times \mathbb{R}^N \to \mathbb{R}^N$ *is a* C^k *Carathéodory function for a* $k \geqslant 2$,

(B) *a neighbourhood* $U_r(\mu_0)$ *is contained in* D *for some* $r > 0$,

(C) *the linearization* $\dot{x} = D_x f(t, \mu_0(t))\, x$ *of* (32) *along* μ_0 *has bounded growth and therefore (Siegmund* [52]*) the dichotomy spectrum consists of* n, $1 \leqslant n \leqslant N$, *compact intervals* $\lambda_i = [a_i, b_i]$, $i = 1, \ldots, n$,

(D) *higher order terms of* f *in* x *along* μ_0 *are uniformly bounded in* t, *i.e., there is a* $M > 0$ *such that*

$$\|D_x^j f(t, \mu_0(t))\| \leqslant M \qquad \text{for a.a.} \quad t \in \mathbb{R} \qquad \text{and all} \qquad j \in \{2, \ldots, k\}.$$

Then (32) *is locally* C^k *equivalent to a differential equation*

$$\boxed{\dot{x} = g(t, x)} \tag{33}$$

with zero reference solution and (33) *is in* normal form, *i.e. it holds that*

(A') $g\colon U_q(0) \to \mathbb{R}^N$ *is a* C^k *Carathéodory function for some* $q > 0$,

(B') *the linearization* $\dot{x} = D_x g(t, 0)\, x$ *of* (33) *along the zero solution has the same dichotomy spectrum as the linearization of* (32) *along* μ_0 *and additionally is block-diagonalized, each block corresponds to a spectral interval* λ_i,

(C') *all nontrivial Taylor components of g of order* 2 *to k are resonant, i.e. for every* $j \in \{1, ..., n\}$ *and* $\ell \in \mathbb{N}_0^n$, $2 \leqslant |\ell| \leqslant k$ *with*

$$\boxed{\lambda_j \cap \sum_{i=1}^{n} \ell_i \lambda_i = \varnothing}$$

we have $D_x^\ell g_j(t, 0) \equiv 0$ *on* $\mathbb{R}$.

Finally we give the normal forms of *scalar equations* and the *Duffing–van der Pol oscillator* under a parametric nonautonomous perturbation. Consider a *scalar equation*

$$\dot{x} = A(t)\, x + F(t, x)$$

with zero reference solution which satisfies the assumptions of the Normal Form Theorem with $N = 1$. The linearization $\dot{x} = A(t)\, x$ along the zero solution has dichotomy spectrum $\Sigma(A) = \lambda = [a, b]$ for some $a, b \in \mathbb{R}$, $a \leqslant b$, and the evolution operator is $\Phi(t, s) = \exp(\int_s^t A(u)\, du)$, $t, s \in \mathbb{R}$. Since $N = 1$, the nonresonance condition for eliminating a Taylor coefficient $\frac{1}{\ell!} D^\ell F(t, 0) \cdot x^\ell$ for a $\ell \in \{2, ..., k\}$ is $\lambda \cap \ell\lambda = \varnothing$ or equivalently one of the two conditions (i) $a > \ell b$ or (ii) $b < \ell a$, hence if $0 \in \lambda$ we have resonance of order ℓ for every $\ell \in \{2, ..., k\}$ and the Normal Form Theorem yields no simplification. On the other hand if $0 \notin \lambda$ then $\lambda \cap \ell\lambda = \varnothing$ for all $\ell > \max\{\frac{a}{b}, \frac{b}{a}\}$ and the ℓth order nonlinearity $\frac{1}{\ell!} D_x^\ell F(t, 0) \cdot x^\ell$ is eliminated by the C^k equivalence $H(t, x) = x + h(t, x)$ where in case of (i) $h(t, x) = \frac{1}{\ell!} \int_t^\infty \Phi(s, t)^{\ell-1} D_x^\ell F(s, 0)\, ds \cdot x^\ell$ and the transformed equation (see step 11 in the proof of the Normal Form Theorem) is

$$\dot{x} = A(t)\, H^{-1}(t, x) + F(t, H^{-1}(t, x)) - \frac{1}{\ell!} D_x^\ell F(t, 0) \cdot [H^{-1}(t, x)]^\ell$$

$$+ A(t)\, h(t, H^{-1}(t, x)) + D_x h(t, H^{-1}(t, x)) \cdot F(t, H^{-1}(t, x)).$$

By construction, this equation has no ℓ-th order nonlinearity and we can proceed to eliminate its $\ell + 1$-st order nonlinearity, and so on up to order k.

Thus if F is a C^∞ Carathéodory function and $\max\{\frac{a}{b}, \frac{b}{a}\} < 2$ we get a formal linearization, but do not know in general wether the limiting transformation is defined on some open tubular neighbourhood of the zero solution. The full Sternberg linearization which would also transform away the infinitely flat terms at $x=0$ is not yet available. However we are able to compute the C^K Normal Form for arbitrary K.

Now we consider the prototypical *Duffing–van der Pol oscillator*

$$\ddot{y} = \alpha y + \beta \dot{y} - y^3 - y^2 \dot{y}. \tag{34}$$

For $\alpha < 0$ fixed and β the bifurcation parameter, the system (34) exhibits a Hopf bifurcation for $\beta = 0$. For $\beta < 0$ fixed and α the bifurcation parameter, it undergoes a pitchfork bifurcation at $\alpha = 0$. Hoping to stimulate the development of a nonautonomous bifurcation theory (see also L. Arnold [2, Ch. 8, 9]) we consider system (34) under the influence of a nonautonomous parametric perturbation. Let α be replaced by $\alpha + \sigma\xi(t)$, where $\xi: \mathbb{R} \to [-1, 1]$ is a bounded measurable function and σ is an intensity parameter. With $x = \binom{y}{\dot{y}}$ the perturbed version of (34) is

$$\dot{x} = \begin{pmatrix} 0 & 1 \\ 0 & \beta \end{pmatrix} x + (\alpha + \sigma\xi(t)) \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} x + \begin{pmatrix} 0 \\ -x_1^3 - x_1^2 x_2 \end{pmatrix}. \tag{35}$$

The linearization of (35) at $x = 0$ is

$$\dot{v} = \begin{pmatrix} 0 & 1 \\ \alpha & \beta \end{pmatrix} v + \sigma\xi(t) \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} v.$$

We choose $\beta = -1$ and treat (α, σ) as a small two-dimensional parameter (the pitchfork scenario if $\xi \equiv 0$). By adding the trivial equations $\dot{\alpha} = 0$, $\dot{\sigma} = 0$ (which we will omit for notational convenience) we can apply the Normal Form Theorem to eliminate all nonresonant terms up to an order $k \geqslant 2$.

As a first step we diagonalize the linear part of (34) at $\alpha = \sigma = 0$ yielding (writing again x for the new coordinate $T^{-1}x$ with $T = \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix}$)

$$\dot{x} = \begin{pmatrix} 0 & 0 \\ 0 & -1 \end{pmatrix} x + [\alpha x_1 - \alpha x_2 + \sigma x_1 \xi(t) - \sigma x_2 \xi(t) - x_1^3 + 2x_1^2 x_2 - x_1 x_2^2] \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

The four components $(x_1, x_2, \alpha, \sigma)$ of the linearized system have one-point spectra $\lambda_1 = \{0\}$, $\lambda_2 = \{-1\}$, $\lambda_3 = \{0\}$ and $\lambda_4 = \{0\}$. The elimination of nonresonant Taylor coefficients with our Normal Form Theorem was done in Siegmund [54] with the computer algebra program MAPLE. The truncated nonautonomous normal form (without the $O((|x_c|+|x_s|)^4+(|\alpha|+|\sigma|)^3)$ terms) is:

$$\dot{x}_c = g^c = [\alpha+\sigma\xi-\alpha^2-\alpha\sigma\xi_1-\sigma^2\xi\xi_1]\, x_c - x_c^3,$$
$$\dot{x}_s = -x_s + g^s = [-1-\alpha-\sigma\xi+\alpha\sigma\xi_2]\, x_s + 2x_c^2 x_s.$$

In these equations the coefficients are

$$\xi_1(t) = \int_t^{\infty} e^{t-\tau}\xi(\tau)\, d\tau, \qquad \xi_2(t) = \int_{-\infty}^{t} e^{-(t-\tau)}\xi(\tau)\, d\tau.$$

REFERENCES

1. R. Abraham, J. E. Marsden, and T. Ratiu, "Manifolds, Tensor Analysis, and Applications," Applied Mathematical Sciences 75, Springer-Verlag, New York, 1983.
2. L. Arnold, "Random Dynamical Systems," Springer-Verlag, Berlin, 1998.
3. V. I. Arnold, "Geometrical Methods in the Theory of Ordinary Differential Equations," A Series of Comprehensive Studies in Mathematics 250, Springer-Verlag, New York, 1983.
4. V. I. Arnold, "Dynamical Systems V," Encyclopaedia of Mathematical Sciences 5, Springer-Verlag, Berlin, 1994.
5. M. Ashkenazi and S.-N. Chow, Normal forms near critical points for differential equations and maps, *IEEE Trans. Circuits Systems* **35** (1988), 850–862.
6. B. Aulbach and T. Wanner, Integral manifolds for carathéodory type differential equations in Banach spaces, *in* "Six Lectures on Dynamical Systems" (B. Aulbach and F. Colonius, Eds.), World Scientific, Singapore, 1996.
7. G. R. Belitskii, Functional equations and conjugacy of local diffeomorphisms of a finite smoothness class, *Funct. Anal. Appl.* **7** (1973), 268–277.
8. G. R. Belitskii, On the conjugacy of germs of nonhyperbolic mappings of finite smoothness class, *Sov. Math. Dokl.* **15** (1974), 1557–1561.
9. G. R. Belitskii, A counterexample to smooth linearization, *Funct. Anal. Appl.* **10** (1976), 135–137.
10. G. R. Belitskii, Equivalence and normal forms of germs of smooth mappings, *Russian Math. Surveys* **33** (1978), 107–177.
11. G. R. Belitskii, Smooth equivalence of germs of vector fields with a single zero eigenvalue or a pair of purely imaginary eigenvalues, *Funct. Anal. Appl.* **20** (1986), 253–259.
12. G. R. Belitskii, Normal forms of germs of C^∞-mappings, *Ukrainian Math. J.* **32** (1981), 300–304.
13. G. R. Belitskii, Obstacles to finitely smooth linearizations, *Funct. Anal. Appl.* **22** (1988), 316–317.
14. P. Bonckaert and F. Dumortier, On a linearization theorem of Sternberg for germs of diffeomorphisms, *Math. Z.* **185** (1984), 115–135.
15. P. Bonckaert, On the continuous dependence of the smooth change of coordinates in parametrized normal form theorems, *J. Differential Equations* **106** (1993), 107–120.
16. A. D. Brjuno, Analytical form of differential equations, *Trans. Moscow Math. Soc.* **25** (1971), 131–288.

17. A. D. Brjuno, The analytic form of differential equations II, *Trans. Moscow Math. Soc.* **26** (1972), 199–239.
18. A. D. Bruno, "Local Methods in Nonlinear Differential Equations," Springer-Verlag, Berlin, 1989.
19. I. U. Bronstein and A. Ya. Kopanskii, "Smooth Invariant Manifolds and Normal Forms," World Scientific, Singapore, 1994.
20. I. U. Bronstein and A. Ya. Kopanskii, Normal forms of vector fields satisfying certain geometric conditions, *Prog. Nonlinear Differential Equations Appl.* **19** (1996), 79–101.
21. C. Carathéodory, "Vorlesungen über reelle Funktionen," third (corrected) ed., Chelsea, New York 1968.
22. K.-T. Chen, Equivalence and decomposition of vector fields about an elementary critical point, *Amer. J. Math.* **85** (1963), 693–722.
23. S.-N. Chow and J. K. Hale, "Methods of Bifurcation Theory," Grundlehren der mathematischen Wissenschaften 251, Springer-Verlag, Berlin, 1996.
24. S.-N. Chow, C. Li, and D. Wang, "Normal Forms and Bifurcation of Planar Vector Fields," Cambridge Univ. Press, Cambridge, UK 1994.
25. S.-N. Chow, K. Lu, and Y.-Q. Shen, Normal form and linearization for quasiperiodic systems, *Trans. Amer. Math. Soc.* **331** (1992), 361–376.
26. F. Colonius and W. Kliemann, "The Dynamics of Control," Birkhäuser, Basel, 2000.
27. W. A. Coppel, "Dichotomies in Stability Theory," Springer Lecture Notes in Mathematics 629, Springer-Verlag, Berlin, 1978.
28. B. D. Craven, "Lebesgue Measure and Integral," Pitman, Boston, 1982.
29. J. D. Crawford, Introduction to bifurcation theory, *Rev. Modern Phys.* **63** (1991), 991–1037.
30. J. D. Dora and L. Stolovitch, Normal forms of differential systems, *in* "Computer Algebra and Differential Equations" (E. Tournier, ed.), London Mathematical Society Lecture Note Series 193, pp. 143–184, Cambridge Univ. Press, Cambridge, UK, 1994.
31. M. S. ElBialy, Linearization of vector fields near resonant hyperbolic rest points, *J. Differential Equations* **118** (1995), 336–370.
32. W. Hahn, "Stability of Motion," Springer-Verlag, Berlin, 1967.
33. Y. Goltser, Continuous normal form of a class of nonautonomous parametrically-perturbed systems and its application, *J. Appl. Math. Mech.* **47** (1984), 14–21.
34. R. R. Gompa, Time-dependent normal form theory and Schrödinger initial value problems, *J. Math. Phys.* **32** (1991), 959–966.
35. Yu. S. Il'yashenko and S. Yu. Yakovenko, Finitely-smooth normal forms of local families of diffeomorphisms and vector fields, *Russian Math. Surveys* **46** (1991), 1–43.
36. A. Katok and B. Hasselblatt, "Introduction to the Modern Theory of Dynamical Systems," pp. 56–104, Cambridge Univ. Press, Cambridge, UK, 1995.
37. A. Ya. Kopanskii, Finitely smooth normal forms of a local family of diffeomorphisms near the center manifold of a fixed point, *Funct. Anal. Appl.* **29** (1995), 210.
38. A. Ya. Kopanskii, The concept of normal form. Topological and formal normal form, Lectures held at the Dynamics Days, 15–17 Dec., 1998, Augsburg.
39. V. V. Kostin, Normal form of nonautonomous systems, *Dopov. Akad. Nauk Ukr. RSR, Ser. A* (1973), 693–696.
40. J. Kurzweil, "Ordinary Differential Equations," Elsevier, Amsterdam, 1986.
41. Y. A. Kuznetsov, "Elements of Applied Bifurcations Theory," Applied Mathematical Sciences 112, Springer-Verlag, New York, 1995.
42. R. C. Robinson, Differentiable conjugacy near compact invariant manifolds, *Bol. Soc. Brasil. Mat.* **2** (1971), 33–44.
43. R. J. Sacker and G. R. Sell, A spectral theory for linear differential systems, *J. Differential Equations* **27** (1978), 320–358.

44. K. Sakamoto, Smooth linearization of vector fields near invariant manifolds, *Hiroshima Math. J.* **24** (1994), 331–355.
45. V. S. Samovol, Linearization of a system of differential equations in the neighborhood of a singular point, *Dokl. Akad. Nauk SSSR* **206** (1972), 1255–1259.
46. V. S. Samovol, A necessary and sufficient condition of smooth linearization of autonomous planar systems in a neighborhood of a critical point, *Math. Notes* **46** (1989), 543–550.
47. V. S. Samovol, Obstacles to local smooth linearization of systems of ordinary differential equations, *Math. Notes* **57** (1995), 662–664.
48. G. R. Sell, "Linearization and Global Dynamics," Proceedings of the International Congress of Mathematicians, Vol. 2, pp. 1283–1296, Warszawa, 1983.
49. G. R. Sell, Obstacles to linearization, *Differential Equations* **20** (1984), 341–345.
50. G. R. Sell, Smooth linearization near a fixed point, *Amer. J. Math.* **107** (1985), 1035–1091.
51. S. Siegmund, "Spektral-Theorie, glatte Faserungen und Normalformen für Differentialgleichungen vom Carathéodory-Typ," Dissertation, Augsburger Mathematisch-Naturwissenschaftliche Schriften 30, Wißner-Verlag, Augsburg, 1999.
52. S. Siegmund, Dichotomy spectrum for nonautonomous differential equations, *J. Dynam. Differential Equations* **14**, No. 1 (2002), in press.
53. S. Siegmund, Reducibility of nonautonomous linear differential equations, *J. London Math. Soc.*, in press
54. S. Siegmund, Normal form of Duffing–van der Pol oscillator under nonautonomous parametric perturbations, *in* "Proceedings of the Third International Conference on Dynamical Systems and Differential Equations, Kennesaw State University, Kennesaw, GA, U.S.A., May 18–21, 2000," *Discrete Contin. Dynam. Systems* (2000), 357–361.
55. H. L. Smith, Normal forms for periodic systems, *J. Math. Anal. Appl.* **113** (1986), 578–600.
56. S. Sternberg, Local C^n transformations of the real line, *Duke Math. J.* **24** (1957), 97–102.
57. S. Sternberg, Local contractions and a theorem of Poincare, *Amer. J. Math.* **79** (1957), 809–824.
58. S. Sternberg, On the structure of local homeomorphisms of euclidian n-space. II, *Amer. J. Math.* **80** (1958), 623–631.
59. S. Sternberg, The structure of local homeomorphisms. III, *Amer. J. Math.* **81** (1959), 578–604.
60. D. Stowe, Linearization in two dimensions, *J. Differential Equations* **63** (1986), 183–226.
61. F. Takens, Partially hyperbolic fixed points, *Topology* **10** (1971), 133–147.
62. F. Takens, Normal forms for certain singularities of vectorfields, *Ann. Inst. Fourier, Grenoble* **23** (1973), 163–195.
63. J.-C. van der Meer, "The Hamiltonian Hopf Bifurcation," Lecture Notes in Mathematics 1160, 1985.
64. W. H. Warner, P. R. Sethna and J. P. Sethna, A generalization of the theory of normal forms, *J. Nonlinear Sci.* (1996), 499–506.
65. S. Wiggins, "Introduction to Applied Nonlinear Dynamical Systems and Chaos," Texts in Applied Mathematics 2, Springer-Verlag, New York, 1990.